

PCT/JPCO/08631

09/890688

06.12.00

JP 00/8631

日本国特許庁

PATENT OFFICE
JAPANESE GOVERNMENT

REC'D 05 FEB 2001	
TPO	PCT

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出願年月日

Date of Application:

2000年 5月30日

出願番号

Application Number:

特願2000-160851

出願人

Applicant(s):

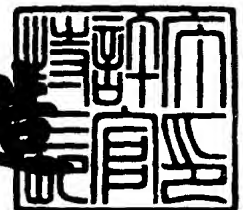
科学技術振興事業団

**PRIORITY
DOCUMENT**
SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH RULE 17.1(a) OR (b)

2001年 1月19日

特許庁長官
Commissioner,
Patent Office

及川耕造



出証番号 出証特2000-3113330

【書類名】 特許願

【整理番号】 NP00238-YS

【提出日】 平成12年 5月30日

【あて先】 特許庁長官 殿

【国際特許分類】 C07K 14/435
C12N 15/12

【発明の名称】 ヒト蛋白質と cDNA [10]

【請求項の数】 7

【発明者】

 【住所又は居所】 神奈川県相模原市若松 3 - 4 6 - 5 0

 【氏名】 加藤 誠志

【発明者】

 【住所又は居所】 神奈川県横浜市戸塚区原宿 1 - 5 3 - 3
 シャーレ原宿 C 1 0 1

 【氏名】 佐伯 美帆呂

【特許出願人】

 【識別番号】 396020800

 【氏名又は名称】 科学技術振興事業団

【代理人】

 【識別番号】 100093230

 【弁理士】

 【氏名又は名称】 西澤 利夫

 【電話番号】 03-5454-7191

【手数料の表示】

 【予納台帳番号】 009911

 【納付金額】 21,000円

【提出物件の目録】

 【物件名】 明細書 1

 【物件名】 図面 1

特 2 0 0 0 - 1 6 0 8 5 1

【物件名】	要約書	1
【プルーフの要否】	要	

【書類名】 明細書

【発明の名称】 ヒト蛋白質と cDNA [10]

【特許請求の範囲】

【請求項1】 配列番号2、4、6、8、10、12、14、16または18のいずれかのアミノ酸配列を有する精製ヒト蛋白質。

【請求項2】 請求項1の蛋白質をコードするDNA断片。

【請求項3】 請求項1の蛋白質をコードするヒトcDNAであって、1、3、5、7、9、11、13、15または17の翻訳領域の塩基配列を有するDNA断片。

【請求項4】 配列番号1、3、5、7、9、11、13、15または17のいずれかの塩基配列からなる請求項3のDNA断片。

【請求項5】 請求項2から4のいずれかのDNA断片をインビトロ翻訳あるいは宿主細胞内で発現しうる発現ベクター。

【請求項6】 請求項5の発現ベクターによる形質転換体であって、請求項1の蛋白質を生産しうる形質転換細胞。

【請求項7】 請求項1記載の蛋白質に対する抗体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

この出願の発明は、精製ヒト蛋白質、この蛋白質をコードしているDNA断片、このDNA断片の発現ベクター、この発現ベクターにより形質転換した各種の細胞、およびこの蛋白質に対する抗体に関するものである。この発明の蛋白質は、医薬品として、あるいはこの蛋白質に対する抗体を作製するための抗原として用いることができる。また、この蛋白質は、細胞内蛋白質ネットワークを解明するための研究試薬として、あるいは低分子医薬と結合する蛋白質をスクリーニングするための蛋白質源として用いることができる。この発明のヒトcDNAは、遺伝子診断用プローブや遺伝子治療用遺伝子源として用いることができる。また、このcDNAがコードしている蛋白質を大量生産するための遺伝子源として用いることができる。これらのDNAをインビトロ翻訳あるいは宿主細胞内で発現

しうる発現ベクターは、この発明の蛋白質をインビトロであるいは各種の宿主細胞内で生産するのに用いることができる。これらの遺伝子を導入して蛋白質を過剰発現させた細胞は、対応するレセプターやリガンドの検出、新しい低分子医薬のスクリーニングなどに利用できる。この発明の蛋白質に対する抗体は、蛋白質を精製するための手段、あるいは細胞内における蛋白質の発現量や局在部位を調べるのに用いられる。

【0002】

【従来の技術】

ヒト蛋白質は、我々の身体を構成している細胞の基本要素である。その中には、（１）細胞の形態を維持したり、細胞内の物質輸送や細胞運動に関わっている細胞骨格蛋白質、（２）細胞内の物質代謝に関与する代謝酵素、（３）エネルギー産生に関わる蛋白質、（４）細胞の増殖・分裂に関わる情報伝達蛋白質、（５）蛋白質の合成に関わる翻訳関連蛋白質、（６）蛋白質の分解に関わるプロテアーゼ関連蛋白質、（７）ゲノムの複製に関与する蛋白質、（８）遺伝子の転写に関与する転写因子、（９）mRNAのスプライシングに関与する核蛋白質などが含まれる。これらの蛋白質は、ヒト細胞の働きを解明する上で重要であるのみならず、医薬品の開発においても有用である。これまで知られている低分子化合物医薬の多くは、細胞内のある特定の蛋白質と結合し、その蛋白質の働きを増強したり、阻止したりすることによって、その薬効を表す。したがって、一揃いのヒト蛋白質を持っていれば、これらの低分子医薬をスクリーニングする際の有力な道具となる。

【0003】

従来、ヒト蛋白質を得るには、ヒト組織や培養細胞をすりつぶした後、各種の分離法を組み合わせることで単一の蛋白質を精製する方法がとられてきた。これまで知られている蛋白質のように、含有量が高く、活性が分かっているものは、従来の方法で容易に単離精製できるが、まだ解析されていない蛋白質の多くは含量が低く、かつその性質によっては単離するのが困難である。また、ヒト組織の多くは入手困難である。したがって、従来のように蛋白質を単離精製する方法では、ヒト蛋白質を全てそろえることは不可能に近い。

【 0 0 0 4 】

一方、ヒト蛋白質の構造情報は、ヒトゲノムDNAに書かれているので、この情報をすべて読み取れば、全ヒト蛋白質の一次構造を推定することができる。ヒトゲノムプロジェクトの目的の一つはここにある。ただ、ゲノム解読の結果得られるのは、DNA配列情報だけであり、蛋白質そのものは得られない。細胞内では、ゲノムの情報はまずmRNAに転写され、mRNAの配列情報を翻訳して蛋白質が合成される。したがって、このmRNAを鋳型にして作製したcDNAが合成できれば、このcDNAを用いて対応する蛋白質も合成することが可能となる。そこで、各種細胞から単離したmRNAを鋳型にして、cDNAを合成し、cDNAの部分塩基配列を決定するいわゆるESTプロジェクトが進行している。

【 0 0 0 5 】

【発明が解決しようとする課題】

蛋白質の取得を目的とする場合、cDNAに要求される必須要件は、蛋白質の翻訳領域を全て含んでいること、いわゆる完全長cDNAであることである。しかしながら、従来法で合成したcDNAは、完全長である割合は低く、得られたものが完全長かどうかを判定することも困難である。すなわち、ESTとして知られているものの多くは蛋白質の翻訳領域の一部のみ含んでいるcDNA断片である。

【 0 0 0 6 】

これに対して、この出願の発明者らは、独自の完全長cDNA合成技術を完成させている(Kato, S. et al., Gene 150:243-250, 1994)。そしてこの技術で合成したヒト完全長cDNAクローンを解析することにより、ヒト蛋白質を完全長cDNAの形で取得することが可能となった。この技術を用いてヒト完全長cDNAをすべてクローン化し、ヒト蛋白質バンクを作製することが望まれている。

【 0 0 0 7 】

また、これまでのヒト疾患に関する研究の結果、ほとんどの病気は何らかの形で遺伝子に異常があるために引き起こされることが明らかになりつつある。これ

らの病気を治療するためには、異常な遺伝子の代わりに正常な遺伝子を導入する遺伝子治療が有望視されている。この際も、ヒトの完全長 cDNA は、遺伝子治療用の遺伝子源として用いることができる。

【0008】

この出願の発明は、以上のとおりの事情に鑑みてなされたものであって、新規の精製ヒト蛋白質、この蛋白質をコードする DNA 断片、この DNA 断片の発現ベクター、この発現ベクターにより形質転換された細胞およびこの蛋白質に対する抗体を提供することを課題としている。

【0009】

【課題を解決するための手段】

この出願は、前記の課題を解決するものとして、以下の(1)～(7)の発明を提供する。

- (1) 配列番号 2、4、6、8、10、12、14、16 または 18 のいずれかのアミノ酸配列を有する精製ヒト蛋白質。
- (2) 前記発明(1)の蛋白質をコードする DNA 断片。
- (3) 前記発明(1)の蛋白質をコードするヒト cDNA であって、1、3、5、7、9、11、13、15 または 17 の翻訳領域の塩基配列を有する DNA 断片。
- (4) 配列番号 1、3、5、7、9、11、13、15 または 17 のいずれかの塩基配列からなる前記発明(3)の DNA 断片。
- (5) 前記発明(2)から(4)のいずれかの DNA 断片をインビトロ翻訳あるいは宿主細胞内で発現しうる発現ベクター。
- (6) 前記発明(5)の発現ベクターによる形質転換体であって、前記発明(1)の蛋白質を生産しうる形質転換細胞。
- (7) 前記発明(1)の蛋白質に対する抗体。

【0010】

【発明の実施の形態】

前記発明(1)の蛋白質は、ヒトの臓器、細胞株などから単離する方法、この出願によって提供されるアミノ酸配列に基づき化学合成によってペプチドを調製する方法、あるいは前記発明(2)～(4)の DNA 断片を用いて組換え DNA 技術で生

産する方法などにより取得することができるが、組換えDNA技術で取得する方法が好ましく用いられる。例えば、前記発明(3)または(4)のDNA断片(cDNA)を有するベクターからインビトロ転写によってRNAを調製し、これを鋳型としてインビトロ翻訳を行なうことによりインビトロで蛋白質を発現できる。また翻訳領域を公知の方法により適当な発現ベクターに組換えることにより、大腸菌、枯草菌等の原核細胞や、酵母、昆虫細胞、哺乳動物細胞、植物細胞等の真核細胞で、DNA断片がコードしている蛋白質を大量に発現させることができる。

【0011】

前記発明(1)の蛋白質をインビトロ翻訳でDNA断片を発現させて生産させる場合には、例えば前記発明(3)または(4)のDNA断片の翻訳領域を、RNAポリメラーゼプロモーターを有するベクターに組換え、プロモーターに対応するRNAポリメラーゼを含む、ウサギ網状赤血球溶解物や小麦胚芽抽出物などのインビトロ翻訳系に添加すれば、前記発明(1)の蛋白質をインビトロで生産することができる。RNAポリメラーゼプロモーターとしては、T7、T3、SP6などが例示できる。これらのRNAポリメラーゼプロモーターを含むベクターとしては、pKA1、pCDM8、pT3/T7 18、pT7/3 19、pBluescript IIなどが例示できる。

【0012】

前記発明(1)の蛋白質を大腸菌などの微生物でDNA断片を発現させて生産させる場合には、微生物中で複製可能なオリジン、プロモーター、リボソーム結合部位、DNAクローニング部位、ターミネーター等を有する発現ベクターに、例えば前記発明(3)または(4)のDNA断片の翻訳領域を組換えた発現ベクターを作成し、この発現ベクターで宿主細胞を形質転換したのち、得られた形質転換体を培養すれば、このDNA断片がコードしている蛋白質を微生物内で大量生産することができる。この際、任意の翻訳領域の前後に開始コドンと停止コドンを付加して発現させれば、任意の領域を含む蛋白質断片を得ることができる。あるいは、他の蛋白質との融合蛋白質として発現させることもできる。この融合蛋白質を適当なプロテアーゼで切断することによってこのcDNAがコードする蛋白質部分のみを取得することもできる。大腸菌用発現ベクターとしては、pUC系、p

Bluescript II、pET発現システム、pGEX発現システムなどが例示できる。

【0013】

前記発明(1)の蛋白質を、真核細胞でDNA断片を発現させて生産させる場合には、例えば前記発明(3)または(4)のDNA断片の翻訳領域を、プロモーター、スプライシング領域、ポリ(A)付加部位等を有する真核細胞用発現ベクターに組換え、真核細胞内に導入すれば、前記発明(1)の蛋白質を真核細胞内で生産することができる。発現ベクターとしては、pKA1、pCDM8、pSVK3、pMSG、pSVL、pBK-CMV、pBK-RSV、EBVベクター、pRS、pYES2などが例示できる。また、pIND/V5-His、pFLAG-CMV-2、pEGFP-N1、pEGFP-C1などを発現ベクターとして用いれば、Hisタグ、FLAGタグ、GFPなど各種タグを付加した融合蛋白質として発現させることもできる。真核細胞としては、サル腎臓細胞COS7、チャイニーズハムスター卵巣細胞CHOなどの哺乳動物培養細胞、出芽酵母、分裂酵母、カイコ細胞、アフリカツメガエル卵細胞などが一般に用いられるが、前記発明(1)の蛋白質を発現できるものであれば、いかなる真核細胞でもよい。発現ベクターを真核細胞に導入するには、電気穿孔法、リン酸カルシウム法、リポソーム法、DEAEデキストラン法など公知の方法を用いることができる。

【0014】

前記発明(1)の蛋白質を原核細胞や真核細胞で発現させたのち、培養物から目的蛋白質を単離精製するためには、公知の分離操作を組み合わせる行うことができる。例えば、尿素などの変性剤や界面活性剤による処理、超音波処理、酵素消化、塩析や溶媒沈殿法、透析、遠心分離、限外濾過、ゲル濾過、SDS-PAGE、等電点電気泳動、イオン交換クロマトグラフィー、疎水性クロマトグラフィー、アフィニティークロマトグラフィー、逆相クロマトグラフィーなどがあげられる。

【0015】

前記発明(1)の蛋白質には、配列番号2、4、6、8、10、12、14、16または18のアミノ酸配列のいかなる部分アミノ酸配列からなるペプチド断片

(5アミノ酸残基以上) も含まれる。これらのペプチド断片は抗体を作製するための抗原として用いることができる。また、前記発明(1)の蛋白質の多くは、翻訳された後、細胞内で各種修飾を受ける。したがって、これらの修飾された蛋白質も前記発明(1)の蛋白質の範囲に含まれる。このような翻訳後修飾としては、N末端メチオニンの脱離、N末端アセチル化、糖鎖付加、細胞内プロテアーゼによる限定分解、ミリスチル化、イソプレニル化、リン酸化などが例示できる。

【0016】

前記発明(2)～(4)のDNA断片には、前記(1)の蛋白質をコードするすべてのDNAが含まれる。このDNA断片は、化学合成による方法、cDNAクローニングによる方法、ヒトゲノムライブラリーをスクリーニングする方法などを用いて取得することができる。

【0017】

前記発明(3)または(4)のDNA断片(cDNA)は、例えばヒト細胞由来cDNAライブラリーからクローン化することができる。cDNAはヒト細胞から抽出したポリ(A)⁺RNAを鋳型として合成する。ヒト細胞としては、人体から手術などによって摘出されたものでも培養細胞でも良い。cDNAは、岡山-Berg法(Okayama, H. and Berg, P., Mol. Cell. Biol. 2:161-170, 1982)、Gubler-Hoffman法(Gubler, U. and Hoffman, J., Gene 25:263-269, 1983)などいかなる方法を用いて合成してもよいが、完全長クローンを効率的に得るためには、実施例にあげたようなキャッピング法(Kato, S. et al., Gene 150:243-250, 1994)を用いることが望ましい。また市販のヒトcDNAライブラリーを用いることもできる。cDNAライブラリーから目的のcDNAをクローン化するには、この出願によって提供される前記発明(3)または(4)のcDNA(配列番号1、3、5、7、9、11、13、15または17)の任意部分の塩基配列に基づいてオリゴヌクレオチドを合成し、これをプローブとして用いて、公知の方法によりコロニーあるいはブラークハイブリダイゼーションによるスクリーニングを行えばよい。また、目的とするcDNA断片の両末端にハイブリダイズするオリゴヌクレオチドを合成し、これをプライマーとして用いて、ヒト細胞から単離したmRNAからRT-PCR法により、前記発明(3)または(4)のcDNA断片を調製

することもできる。

【0018】

前記発明(3)のDNA断片は、配列番号1、3、5、7、9、11、13、15または17の翻訳領域(Open Reading Frame: ORF)の塩基配列を有するcDNAであり、前記発明(4)のDNA断片は、配列番号1、3、5、7、9、11、13、15または17のいずれかの塩基配列からなるcDNAである。それぞれのクローン番号(HP番号)、cDNAクローンが得られた細胞、cDNAの全塩基数、コードしている蛋白質のアミノ酸残基数をそれぞれ表1にまとめて示した。

【0019】

【表1】

表1

配列番号	HP番号	細胞	塩基数	アミノ酸 残基数
1、2	HP01124	肝臓	1664	341
3、4	HP02241	胃癌	835	216
5、6	HP10101	HT-1080	2465	396
7、8	HP10370	KB	3600	451
9、10	HP10427	胃癌	442	113
11、12	HP10438	胃癌	726	222
13、14	HP10502	HT-1080	1120	278
15、16	HP10516	胃癌	747	221
17、18	HP10580	胃癌	1441	441

【0020】

なお、配列番号1、3、5、7、9、11、13、15または17のいずれかの塩基配列に基づいて合成したオリゴヌクレオチドプローブを用いて、表1に示したヒト細胞株やヒト組織から作製したcDNAライブラリーをスクリーニングすることにより、前記発明(3)および(4)のcDNAと同一のクローンを容易に得ることができる。

【0021】

また、一般にヒト遺伝子は個体差による多型が頻繁に認められる。従って配列番号1、3、5、7、9、11、13、15または17において、1または複数個のヌクレオチドの付加、欠失および／または他のヌクレオチドによる置換がなされているcDNAもこの発明の範囲に含まれる。

【0022】

同様に、これらの変更によって生じる1または複数個のアミノ酸の付加、欠失および／または他のアミノ酸による置換がなされている蛋白質も、配列番号2、4、6、8、10、12、14、16または18のアミノ酸配列を有するそれぞれの蛋白質の活性を有する限り、この発明の範囲に含まれる。

【0023】

前記発明(3)および(4)のDNA断片には、配列番号1、3、5、7、9、11、13、15または17の塩基配列のいかなる部分塩基配列からなるDNA断片(10bp以上)も含まれる。また、センス鎖およびアンチセンス鎖からなるDNA断片もこの範囲に含まれる。これらのDNA断片は遺伝子診断用のプローブとして用いることができる。

【0024】

前記発明(7)の抗体は、前記発明(1)の蛋白質を抗原として用いて動物を免疫した後、血清から得ることが出来る。抗原としては配列番号2、4、6、8、10、12、14、16または18のアミノ酸配列に基づいて化学合成したペプチドや、真核細胞や原核細胞で発現させた蛋白質を用いることが出来る。あるいは、上記の真核細胞用発現ベクターを注射や遺伝子銃によって、動物の筋肉や皮膚に導入した後、血清を採取することによって作製することができる(例えば、特開平7-313187号公報記載の方法)。動物としては、マウス、ラット、ウサギ、ヤギ、ニワトリなどが用いられる。免疫した動物の脾臓から採取したB細胞をミエローマと融合させてハイブリドーマを作製すれば、前記発明(1)の蛋白質に対するモノクローナル抗体を産生することができる。

【0025】

【実施例】

次に実施例を示してこの出願の発明をさらに詳細かつ具体的に説明するが、こ

の出願の発明は以下の例によって限定されるものではない。なお、以下の実施例において、DNAの組換えに関する基本的な操作および酵素反応は、文献（"Molecular Cloning. A Laboratory Manual", Cold Spring Harbor Laboratory, 1989）の記載に方従った。制限酵素および各種修飾酵素は特に記載の無い場合は宝酒造社製のものを用いた。各酵素反応の緩衝液組成、並びに反応条件は付属の説明書に従った。cDNA合成は文献（Kato, S. et al., Gene 150:243-250, 1994）の記載に従った。

実施例 1 : cDNA クローニング

cDNAライブラリーとして、ヒト完全長cDNAライブラリー（WO97/33993、WO98/11217、WO98/21328記載）を用いた。個々のライブラリーから完全長cDNAクローンを選択し、その全塩基配列決定を行った。得られたクローン（A）～（I）の詳細は以下のとおりである。

（A） HP01124

ヒト肝臓cDNAライブラリーから得られたクローンHP01124のcDNAインサートの全塩基配列を決定したところ、105bpの5' 非翻訳領域、1026bpのORF、533bpの3' 非翻訳領域からなる構造を有していた（配列番号1）。ORFは341アミノ酸残基（配列番号2）からなる蛋白質をコードしており、インビトロ翻訳の結果、ORFから予想される分子量37,786とほぼ同じ37kDaの翻訳産物が生成した（実施例2）。この蛋白質とGFPとの融合蛋白質は、核に発現が認められた（実施例4）。

【0026】

この蛋白質のアミノ酸配列を用いてプロテインデータベースを検索したところ、ヒトアシルCoA結合蛋白質（アクセション番号P07108）と類似性を有していた。図1に、クローン（A）がコードするヒト蛋白質と、ヒトアシルCoA結合蛋白質のアミノ酸配列の比較を示す。－はギャップを、*はこの発明の蛋白質と同一アミノ酸残基を、. はこの発明の蛋白質と類似アミノ酸残基をそれぞれ表す。全領域にわたって、43.0%の相同性を有していた。

【0027】

クローン（A）cDNAの塩基配列を用いてGenBankを検索したところ

、E S Tの中に90%以上の相同性を有するもの（例えば、アクセシオン番号N 4 1 5 4 2）が登録されていたが、部分配列なのでクローン（A）がコードする蛋白質と同じ蛋白質をコードしているかどうかは判定できない。

（B） H P 0 2 2 4 1

ヒト胃癌cDNAライブラリーから得られたクローンH P 0 2 2 4 1のcDNAインサートの全塩基配列を決定したところ、89bpの5' 非翻訳領域、651bpのORF、95bpの3' 非翻訳領域からなる構造を有していた（配列番号3）。ORFは216アミノ酸残基（配列番号4）からなる蛋白質をコードしており、インビトロ翻訳の結果、ORFから予想される分子量24, 899より大きい30kDaの翻訳産物が生成した（実施例2）。この蛋白質とGFPとの融合蛋白質は、細胞全体に一部凝集塊となって発現が認められた（実施例4）。

【0028】

この蛋白質のアミノ酸配列を用いてプロテインデータベースを検索したところ、アフリカツメガエルリボソーム蛋白質L24様蛋白質（アクセシオン番号C A B 4 0 5 5 4）と類似性を有していた。図2に、クローン（B）がコードするヒト蛋白質と、アフリカツメガエルリボソーム蛋白質L24様蛋白質のアミノ酸配列の比較を示す。－はギャップを、＊はこの発明の蛋白質と同一アミノ酸残基を、. はこの発明の蛋白質と類似アミノ酸残基をそれぞれ表す。N末端側208アミノ酸残基にわたって、69.7%の相同性を有していた。

【0029】

クローン（B）cDNAの塩基配列を用いてG e n B a n kを検索したところ、E S Tの中に、90%以上の相同性を有するもの（例えば、アクセシオン番号A L 0 3 8 4 9 3）が登録されていたが、部分配列なのでクローン（B）がコードする蛋白質と同じ蛋白質をコードしているかどうかは判定できない。

（C） H P 1 0 1 0 1

ヒトフィブロサルコーマ細胞株H T - 1 0 8 0 cDNAライブラリーから得られたクローンH P 1 0 1 0 1のcDNAインサートの全塩基配列を決定したところ、70bpの5' 非翻訳領域、1191bpのORF、1204bpの3' 非翻訳領域からなる構造を有していた（配列番号5）。ORFは396アミノ酸残

基（配列番号 6）からなる蛋白質をコードしていた。インビトロ翻訳の結果、ORF から予想される分子量 45,750 より大きい 54 kDa の翻訳産物が生成した（実施例 2）。この蛋白質と GFP との融合蛋白質は、核内に粒状の発現が認められた（実施例 4）。

【0030】

この蛋白質のアミノ酸配列を用いてプロテインデータベースを検索したところ、線虫仮想蛋白質 C32E8.5（アクセシオン番号 AAB42323）と類似性を有していた。図 3 に、クローン（C）がコードするヒト蛋白質と、線虫仮想蛋白質 C32E8.5 のアミノ酸配列の比較を示す。－はギャップを、*はこの発明の蛋白質と同一アミノ酸残基を、. はこの発明の蛋白質と類似アミノ酸残基をそれぞれ表す。C 末端側 307 アミノ酸残基にわたって、46.9% の相同性を有していた。

【0031】

クローン（C）cDNA の塩基配列を用いて GenBank を検索したところ、EST の中に、90% 以上の相同性を有するもの（例えば、アクセシオン番号 AA460870）が登録されていたが、部分配列なのでクローン（C）がコードする蛋白質と同じ蛋白質をコードしているかどうかは判定できない。

(D) HP10370

ヒト類表皮癌細胞株 KB cDNA ライブラリーから得られたクローン HP10370 の cDNA インサートの全塩基配列を決定したところ、148bp の 5' 非翻訳領域、1356bp の ORF、2096bp の 3' 非翻訳領域からなる構造を有していた（配列番号 7）。ORF は 451 アミノ酸残基（配列番号 8）からなる蛋白質をコードしていた。この蛋白質と GFP との融合蛋白質は、細胞全体に発現が認められた（実施例 4）。

【0032】

この蛋白質のアミノ酸配列を用いてプロテインデータベースを検索したところ、ショウジョウバエ仮想蛋白質 CG11534（アクセシオン番号 AAF49957）と類似性を有していた。図 4 に、クローン（D）がコードするヒト蛋白質と、ショウジョウバエ仮想蛋白質 CG11534 のアミノ酸配列の比較を示す。

－はギャップを、＊はこの発明の蛋白質と同一アミノ酸残基を、．はこの発明の蛋白質と類似アミノ酸残基をそれぞれ表す。C末端側 3 8 2 アミノ酸残基にわたって、3 6 . 9 %の相同性を有していた。

【 0 0 3 3 】

クローン (D) cDNAの塩基配列を用いてG e n B a n kを検索したところ、E S Tの中に、9 0 %以上の相同性を有するもの（例えば、アクセシオン番号 A A 0 3 5 3 2 2）が登録されていたが、部分配列なのでクローン (D) がコードする蛋白質と同じ蛋白質をコードしているかどうかは判定できない。

(E) H P 1 0 4 2 7

ヒト胃癌cDNAライブラリーから得られたクローンH P 1 0 4 2 7のc D N Aインサートの全塩基配列を決定したところ、1 1 b pの5' 非翻訳領域、3 4 2 b pのO R F、8 9 b pの3' 非翻訳領域からなる構造を有していた（配列番号9）。O R Fは1 1 3アミノ酸残基（配列番号1 0）からなる蛋白質をコードしていた。この蛋白質とG F Pとの融合蛋白質は、ゴルジ体に局在が認められた（実施例4）。

【 0 0 3 4 】

この蛋白質のアミノ酸配列を用いてプロテインデータベースを検索したところ、線虫仮想蛋白質Y 1 0 6 G 6 H. 8（アクセシオン番号C A B 6 3 3 8）と類似性を有していた。図5に、クローン (E) がコードするヒト蛋白質と、線虫仮想蛋白質Y 1 0 6 G 6 H. 8のアミノ酸配列の比較を示す。－はギャップを、＊はこの発明の蛋白質と同一アミノ酸残基を、．はこの発明の蛋白質と類似アミノ酸残基をそれぞれ表す。全領域にわたって、3 6 . 9 %の相同性を有していた。

【 0 0 3 5 】

クローン (E) cDNAの塩基配列を用いてG e n B a n kを検索したところ、E S Tの中に、9 0 %以上の相同性を有するもの（例えば、アクセシオン番号 R 7 6 1 7 8）が登録されていたが、部分配列なのでクローン (E) がコードする蛋白質と同じ蛋白質をコードしているかどうかは判定できない。

(F) H P 1 0 4 3 8

ヒト胃癌cDNAライブラリーから得られたクローンH P 1 0 4 3 8のc D N

A インサートの全塩基配列を決定したところ、11bpの5' 非翻訳領域、669bpのORF、46bpの3' 非翻訳領域からなる構造を有していた（配列番号11）。ORFは222アミノ酸残基（配列番号12）からなる蛋白質をコードしており、インビトロ翻訳の結果、ORFから予想される分子量25,384よりやや大きい28kDaの翻訳産物が生成した（実施例2）。この蛋白質とGFPとの融合蛋白質は、核に発現が認められた（実施例4）。

【0036】

クローン（F）cDNAの塩基配列を用いてGenBankを検索したところ、ESTの中に、90%以上の相同性を有するもの（例えば、アクセッション番号AA088470）が登録されていたが、部分配列なのでクローン（F）がコードする蛋白質と同じ蛋白質をコードしているかどうかは判定できない。

（G） HP10502

ヒトフィブロサルコーマ細胞株HT-1080cDNAライブラリーから得られたクローンHP10502のcDNAインサートの全塩基配列を決定したところ、207bpの5' 非翻訳領域、837bpのORF、76bpの3' 非翻訳領域からなる構造を有していた（配列番号13）。ORFは278アミノ酸残基（配列番号14）からなる蛋白質をコードしていた。この蛋白質とGFPとの融合蛋白質は、核に発現が認められた（実施例4）。

【0037】

クローン（G）cDNAの塩基配列を用いてGenBankを検索したところ、ESTの中に、90%以上の相同性を有するもの（例えば、アクセッション番号AA648423）が登録されていたが、部分配列なのでクローン（G）がコードする蛋白質と同じ蛋白質をコードしているかどうかは判定できない。

（H） HP10516

ヒト胃癌cDNAライブラリーから得られたクローンHP10516のcDNAインサートの全塩基配列を決定したところ、26bpの5' 非翻訳領域、666bpのORF（配列番号15）、55bpの3' 非翻訳領域からなる構造を有していた（配列番号16）。ORFは221アミノ酸残基（配列番号8）からなる蛋白質をコードしていた。この蛋白質とGFPとの融合蛋白質は、細胞全体に

発現が認められた（実施例 4）。

【 0 0 3 8 】

この蛋白質のアミノ酸配列を用いてプロテインデータベースを検索したところ、ショウジョウバエ仮想蛋白質 CG 1 4 1 3 0（アクセシオン番号 A A F 5 0 0 0 5）と類似性を有していた。図 6 に、クローン（H）がコードするヒト蛋白質と、ショウジョウバエ仮想蛋白質 CG 1 4 1 3 0 のアミノ酸配列の比較を示す。－はギャップを、＊はこの発明の蛋白質と同一アミノ酸残基を、. はこの発明の蛋白質と類似アミノ酸残基をそれぞれ表す。全領域にわたって、3 6 . 9 % の相同性を有していた。

【 0 0 3 9 】

クローン（H）c D N A の塩基配列を用いて G e n B a n k を検索したところ、E S T の中に、9 0 % 以上の相同性を有するもの（例えば、アクセシオン番号 A W 2 4 5 5 5 6）が登録されていたが、部分配列なのでクローン（H）がコードする蛋白質と同じ蛋白質をコードしているかどうかは判定できない。

（ I ） H P 1 0 5 8 0

ヒト胃癌 c D N A ライブラリーから得られたクローン H P 1 0 5 8 0 の c D N A インサートの全塩基配列を決定したところ、9 4 b p の 5' 非翻訳領域、1 3 2 6 b p の O R F（配列番号 1 7）、2 1 b p の 3' 非翻訳領域からなる構造を有していた（配列番号 1 8）。O R F は 4 4 1 アミノ酸残基（配列番号 9）からなる蛋白質をコードしていた。この蛋白質と G F P との融合蛋白質は、細胞質に凝集塊の発現が認められた（実施例 4）。

【 0 0 4 0 】

この蛋白質のアミノ酸配列を用いてプロテインデータベースを検索したところ、ショウジョウバエ仮想蛋白質 CG 5 4 6 9（アクセシオン番号 A A F 5 0 0 0 5）と類似性を有していた。図 7 に、クローン（I）がコードするヒト蛋白質と、ショウジョウバエ仮想蛋白質 CG 5 4 6 9 のアミノ酸配列の比較を示す。－はギャップを、＊はこの発明の蛋白質と同一アミノ酸残基を、. はこの発明の蛋白質と類似アミノ酸残基をそれぞれ表す。全領域にわたって、3 5 . 0 % の相同性を有していた。

【0041】

クローン (I) cDNA の塩基配列を用いて GenBank を検索したところ、EST の中に、90% 以上の相同性を有するもの (例えば、アクセシオン番号 A1188741) が登録されていたが、部分配列なのでクローン (I) がコードする蛋白質と同じ蛋白質をコードしているかどうかは判定できない。

実施例 2 : インビトロ翻訳による蛋白質合成

実施例 1 で単離した cDNA を有するプラスミドベクターを用いて、T_NT ウサギ網状赤血球溶解物キット (プロメガ社製) によるインビトロ転写/翻訳を行なった。この際 [³⁵S] メチオニンを添加し、発現産物をラジオアイソトープでラベルした。いずれの反応もキットに付属のプロトコールに従って行なった。

【0042】

具体的な方法は次のとおりである。プラスミド 2 μ g を、T_NT ウサギ網状赤血球溶解物 12.5 μ l、緩衝液 (キットに付属) 0.5 μ l、アミノ酸混合液 (メチオニンを含まない) 2 μ l、[³⁵S] メチオニン (アマーシャム社) 2 μ l (0.37 MBq/ μ l)、T7 RNA ポリメラーゼ 0.5 μ l、RNasin 20 U を含む総量 25 μ l の反応液中で 30℃、90 分間反応させた。反応液 3 μ l に SDS サンプリングバッファー (125 mM トリス塩酸緩衝液、pH 6.8、120 mM 2-メルカプトエタノール、2% SDS 溶液、0.025% ブロモフェノールブルー、20% グリセロール) 2 μ l を加え、95℃ 3 分間加熱処理した後、SDS-ポリアクリルアミドゲル電気泳動にかけた。オートラジオグラフィーを行ない、翻訳産物の分子量を求めた。

実施例 3 : COS7 細胞による発現

実施例 1 で単離した cDNA を保有する発現ベクターによって形質転換した大腸菌を 100 μ g/ml アンピシリン含有 2xYT 培地 2 ml 中で 37℃ 2 時間培養した後、ヘルパーファージ M13KO7 (50 μ l) を添加し、37℃ で一晩培養した。遠心によって分離した上澄からポリエチレングリコール沈殿によって一本鎖ファージ粒子を得た。これを 100 μ l の 1 mM トリス-0.1 mM EDTA、pH 8 (TE) に懸濁した。

【0043】

サル腎臓由来培養細胞COS7は、10%ウシ胎児血清を含むダルベッコ改変イーグル (DMEM) 培地中、5%CO₂存在下、37℃で培養した。1×10⁵個のCOS7細胞を6穴プレート (ヌンク社、穴の直径3cm) に植え、5%CO₂存在下、37℃で22時間培養した。培地除去後、リン酸緩衝液で細胞表面を洗浄し、さらに50mMトリス塩酸 (pH7.5) を含むDMEM (TDMEM) で再度洗浄した。この細胞に一本鎖ファージ懸濁液1μl、DMEM培地0.6ml、TRANSFECTAMTM (IBF社) 3μlを懸濁したものを添加し、5%CO₂存在下、37℃で3時間培養した。サンプル液を除去後、TDMEMで細胞表面を洗浄し、10%ウシ胎児血清含有DMEMを1穴あたり2ml加え、5%CO₂存在下、37℃にて2日間培養した。培地を [³⁵S] システインあるいは [³⁵S] メチオニンを含む培地に交換した後、1時間培養した。遠心分離によって、培地と細胞を分けたあと、細胞画分の蛋白質をSDS-PAGEにかけた。

実施例4：緑色蛍光蛋白質 (GFP) 融合蛋白質の発現

EcoRI 認識部位を付加した翻訳開始コドンから始まる26merのセンスプライマーとBamHI 認識部位を付加した停止コドンまでを含む26merのアンチセンスプライマーを用い、目的蛋白質をコードするcDNAを鋳型としてPCRにより翻訳領域を増幅した。PCR産物をEcoRIとBamHIで消化し、GFP融合蛋白質発現用ベクターpEGFP-N1 (Clontec社製) のEcoRI-BamHI部位に挿入した。塩基配列を確認した後、得られた融合遺伝子発現ベクターを実施例3に記載の方法によりCOS7細胞にトランスフェクトした。蛍光顕微鏡により緑色蛍光の分布を観察し、目的蛋白質の局在部位を調べた。

実施例5：抗体の作製

EcoRI 認識部位を付加した翻訳開始コドンから始まる26merのセンスプライマーとSalI 認識配列を付加した停止コドンまでを含む26merのアンチセンスプライマーを用い、各cDNAを鋳型としてPCRにより翻訳領域を増幅した。PCR産物をEcoRIとSalIで消化し、pGEX-5X-1 (ファルマシア社製) のEcoRIとSalI部位に挿入した。塩基配列を確認した後

、宿主大腸菌 JM109 の形質転換を行った。LB 培地中で 37℃、5 時間培養し、IPTG を最終濃度が 0.4 mM になるように加え、さらに 37℃ で 4 時間培養した。菌体を遠心により分離し、溶解溶液 (50 mM Tris-HCl pH 7.5、1 mM EDTA、0.2 mM PMSF) に溶かし、一度 -80℃ で凍結させ融解させた後、超音波破碎を行った。10,000 × g で 30 分遠心し、上清にグルタチオンセファロース 4 B を加え、4℃ で 1 時間インキュベートした。ビーズを十分洗浄した後、溶出溶液 (50 mM Tris-HCl pH 7.5、50 mM グルタチオン) で融合蛋白質を溶出した。得られた融合蛋白質を抗原として家兔に常法により免疫を行い抗血清を得た。抗血清はまず、40% 飽和硫酸沈殿画分を GST アフィニティーカラムにより GST 抗体を除いた。素通り画分をさらに GST 融合蛋白質の抗原カラムにより精製した。

【0044】

【発明の効果】

以上詳しく説明したとおり、この出願によって、新規な精製ヒト蛋白質、これらの蛋白質をコードしている DNA 断片、この DNA 断片の発現ベクター、この発現ベクターによる形質転換細胞、およびこの蛋白質に対する抗体が提供される。この出願によって提供される蛋白質は、いずれも細胞内で機能している蛋白質と考えられるため、細胞内ターゲット蛋白質として、対応するレセプターやリガンドの検出、新しい低分子医薬のスクリーニングなどに利用できる。またこの蛋白質に対する抗体を作製するための抗原として用いることができる。この出願によって提供される DNA 断片は、遺伝子診断用プローブや遺伝子治療用遺伝子源として用いることができる。また、この DNA 断片を用いることにより、この蛋白質を大量に発現することができる。これら遺伝子を導入してこの蛋白質を発現させた細胞は、この蛋白質の修飾型を得るのに利用できる。この出願によって提供される抗体は、この発明の蛋白質の検出、定量、精製などに利用できる。

【0045】

【配列表】

SEQUENCE LISTING

<110> Japan Science and Technology Corporation

<120> Human Proteins and cDNAs thereof (10)

<130> NP00238-YS

<140>

<141>

<160> 18

<170> PatentIn Ver. 2.1

<210> 1

<211> 1664

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (106)..(1131)

<400> 1

tctggcactg aggccctcgc cacctgcctc gccacctggc gaccctgacc ccaccacact 60

gccttgagag tcgctcaaaa gtagggcccc agggctcgca gcagc atg ggc acc gag 117

Met Gly Thr Glu

1

aaa gaa agc cca gag ccc gac tgc cag aaa cag ttc cag gct gca gtg 165
 Lys Glu Ser Pro Glu Pro Asp Cys Gln Lys Gln Phe Gln Ala Ala Val
 5 10 15 20

agc gtc atc cag aac ctg ccc aag aac ggt tct tac cgc ccc tcc tat 213
 Ser Val Ile Gln Asn Leu Pro Lys Asn Gly Ser Tyr Arg Pro Ser Tyr
 25 30 35

gaa gag atg ctg cga ttc tac agt tac tac aag cag gcc acc atg ggg 261
 Glu Glu Met Leu Arg Phe Tyr Ser Tyr Tyr Lys Gln Ala Thr Met Gly
 40 45 50

ccc tgc ctg gtc ccc cgg ccc ggg ttc tgg gac ccc att gga cga tat 309
 Pro Cys Leu Val Pro Arg Pro Gly Phe Trp Asp Pro Ile Gly Arg Tyr
 55 60 65

aag tgg gac gcc tgg aac agt ctg ggc aag atg agc agg gag gag gcc 357
 Lys Trp Asp Ala Trp Asn Ser Leu Gly Lys Met Ser Arg Glu Glu Ala
 70 75 80

atg tct gcc tac atc act gaa atg aaa ctg gtg gca cag aag gtg atc 405
 Met Ser Ala Tyr Ile Thr Glu Met Lys Leu Val Ala Gln Lys Val Ile
 85 90 95 100

gac aca gtg ccc ctg ggt gag gtg gca gag gac atg ttt ggt tac ttc 453
 Asp Thr Val Pro Leu Gly Glu Val Ala Glu Asp Met Phe Gly Tyr Phe
 105 110 115

gag ccc ctg tac cag gtg atc cct gac atg ccg agg ccc cca gag acc 501

Glu Pro Leu Tyr Gln Val Ile Pro Asp Met Pro Arg Pro Pro Glu Thr

120

125

130

ttc ctg aga agg gtc aca ggt tgg aaa gag cag gtt gtg aat gga gat 549

Phe Leu Arg Arg Val Thr Gly Trp Lys Glu Gln Val Val Asn Gly Asp

135

140

145

gtt ggg gct gtt tca gag cct ccc tgc ctc ccc aag gaa ccg gca ccc 597

Val Gly Ala Val Ser Glu Pro Pro Cys Leu Pro Lys Glu Pro Ala Pro

150

155

160

cca agc cca gct tcc ctc tgg gca gta act cta cca acc cct cca cag 645

Pro Ser Pro Ala Ser Leu Trp Ala Val Thr Leu Pro Thr Pro Pro Gln

165

170

175

180

agt ccc att cac cca ggg acc tgg act ccg agg ttt tct gtg att ccc 693

Ser Pro Ile His Pro Gly Thr Trp Thr Pro Arg Phe Ser Val Ile Pro

185

190

195

tgg agc agc tgg agc ctg agc tgg ttt gga cag agc agc ggg cag cat 741

Trp Ser Ser Trp Ser Leu Ser Trp Phe Gly Gln Ser Ser Gly Gln His

200

205

210

ctg gag gaa agc gtg atc cca gga aca gcc ccg tgc ccc cca caa aga 789

Leu Glu Glu Ser Val Ile Pro Gly Thr Ala Pro Cys Pro Pro Gln Arg

215

220

225

aag agg ggt tgc ggg gca gcc cgc cgg ggc ccc agg agt tgg acg tgt 837

Lys Arg Gly Cys Gly Ala Ala Arg Arg Gly Pr Arg Ser Trp Thr Cys

230	235	240	
ggc tgc tgg gga cag ttc gag cac tac agg aga gca tgc agg agg tgc	885		
Gly Cys Trp Gly Gln Phe Glu His Tyr Arg Arg Ala Cys Arg Arg Cys			
245	250	255	260
agg cga ggg tgc aga gcc tgg aga gca tgc ccc ggc ccc ctg agc aga	933		
Arg Arg Gly Cys Arg Ala Trp Arg Ala Cys Pro Gly Pro Leu Ser Arg			
	265	270	275
ggc cgc agc cca ggc cca gtg ctc ggc cat ggc ccc ttg ggc tcc cgg	981		
Gly Arg Ser Pro Gly Pro Val Leu Gly His Gly Pro Leu Gly Ser Arg			
	280	285	290
ggc ccg cgc tgc tct tct tcc tcc tgt ggc cct tcg tcg tcc agt ggc	1029		
Gly Pro Arg Cys Ser Ser Ser Ser Cys Gly Pro Ser Ser Ser Ser Gly			
	295	300	305
tct tcc gaa tgt ttc gga ccc aaa aga ggt gac tgt cag tgg agg ggt	1077		
Ser Ser Glu Cys Phe Gly Pro Lys Arg Gly Asp Cys Gln Trp Arg Gly			
	310	315	320
ctc tgc agc caa ctg aga cta tct tgc tgt gcc ctg agc ctt cct agg	1125		
Leu Cys Ser Gln Leu Arg Leu Ser Cys Cys Ala Leu Ser Leu Pro Arg			
	325	330	335
ggt tag aagaacagca ttcaaaattc cccgtcctgt cagtgtttgc cttcgacct	1181		
Val			

cctccccctaa agcagcgcgagg ggggcaaata agacccccacc cctccctgca gcttcacagg 1241

gacgcttcct tccctccccg caaccacccc aggcctcccct gggaggctgc agttgtggta 1301

cacgtccccg gtgctgggtt ggccgtgact cgggggcggg gcgatcgggt ctcagcccct 1361

gccttcccca gtctctgggt caccgaatt tccccacccc tgcttctccc cgaggaggtt 1421

gagctcttga gcaagttggg acttgggccg gggcctggaa gaatgattgg ctgggaggcc 1481

gcgggagggg ggccaggagg cccggaccag ttgggaggag tgagcaggcc ccgggggagg 1541

gggatgagcg cagtttgctc gctttcctcc cctgccggcc ccctccgccc ccacacacac 1601

tcgggacgtc ttcatigaag attcacttac aaaggaatgt ttcactaaat aaaagaaaac 1661

cag 1664

<210> 2

<211> 341

<212> PRT

<213> Homo sapiens

<400> 2

Met Gly Thr Glu Lys Glu Ser Pro Glu Pro Asp Cys Gln Lys Gln Phe

1

5

10

15

Gln Ala Ala Val Ser Val Ile Gln Asn Leu Pro Lys Asn Gly Ser Tyr

20

25

30

Arg Pro Ser Tyr Glu Glu Met Leu Arg Phe Tyr Ser Tyr Tyr Lys Gln			
35	40	45	
Ala Thr Met Gly Pro Cys Leu Val Pro Arg Pro Gly Phe Trp Asp Pro			
50	55	60	
Ile Gly Arg Tyr Lys Trp Asp Ala Trp Asn Ser Leu Gly Lys Met Ser			
65	70	75	80
Arg Glu Glu Ala Met Ser Ala Tyr Ile Thr Glu Met Lys Leu Val Ala			
85	90	95	
Gln Lys Val Ile Asp Thr Val Pro Leu Gly Glu Val Ala Glu Asp Met			
100	105	110	
Phe Gly Tyr Phe Glu Pro Leu Tyr Gln Val Ile Pro Asp Met Pro Arg			
115	120	125	
Pro Pro Glu Thr Phe Leu Arg Arg Val Thr Gly Trp Lys Glu Gln Val			
130	135	140	
Val Asn Gly Asp Val Gly Ala Val Ser Glu Pro Pro Cys Leu Pro Lys			
145	150	155	160
Glu Pro Ala Pro Pro Ser Pro Ala Ser Leu Trp Ala Val Thr Leu Pro			
165	170	175	
Thr Pro Pro Gln Ser Pro Ile His Pro Gly Thr Trp Thr Pro Arg Phe			
180	185	190	
Ser Val Ile Pro Trp Ser Ser Trp Ser Leu Ser Trp Phe Gly Gln Ser			
195	200	205	
Ser Gly Gln His Leu Glu Glu Ser Val Ile Pro Gly Thr Ala Pro Cys			
210	215	220	
Pro Pro Gln Arg Lys Arg Gly Cys Gly Ala Ala Arg Arg Gly Pro Arg			
225	230	235	240
Ser Trp Thr Cys Gly Cys Trp Gly Gln Phe Glu His Tyr Arg Arg Ala			
245	250	255	
Cys Arg Arg Cys Arg Arg Gly Cys Arg Ala Trp Arg Ala Cys Pro Gly			

260 265 270
 Pro Leu Ser Arg Gly Arg Ser Pro Gly Pro Val Leu Gly His Gly Pro
 275 280 285
 Leu Gly Ser Arg Gly Pro Arg Cys Ser Ser Ser Ser Cys Gly Pro Ser
 290 295 300
 Ser Ser Ser Gly Ser Ser Glu Cys Phe Gly Pro Lys Arg Gly Asp Cys
 305 310 315 320
 Gln Trp Arg Gly Leu Cys Ser Gln Leu Arg Leu Ser Cys Cys Ala Leu
 325 330 335
 Ser Leu Pro Arg Val
 340

<210> 3

<211> 835

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (90)..(740)

<400> 3

aaaaatccga agtgccgcgg aaagtgagg gctgacaagg aaggtttcga gcgttttgct 60

ggcaaaggga tttcttataa cctccaggc atg cgt ctt tct gcc ctg ctg gcc 113

Met Arg Leu Ser Ala Leu Leu Ala

1

5

ttg gca tcc aag gtc act ctg ccc ccc cat tac cgc tat ggg atg agc 161
Leu Ala Ser Lys Val Thr Leu Pro Pro His Tyr Arg Tyr Gly Met Ser

10

15

20

ccc cca ggc tct gtt gca gac aag agg aag aac ccc cca tgg atc agg 209
Pro Pro Gly Ser Val Ala Asp Lys Arg Lys Asn Pro Pro Trp Ile Arg

25

30

35

40

cgg cgc cca gtg gtt gtg gaa ccc atc tct gat gaa gac tgg tat ctg 257
Arg Arg Pro Val Val Val Glu Pro Ile Ser Asp Glu Asp Trp Tyr Leu

45

50

55

ttc tgt ggg gac acg gtg gag atc cta gaa ggc aag gat gcc ggg aag 305
Phe Cys Gly Asp Thr Val Glu Ile Leu Glu Gly Lys Asp Ala Gly Lys

60

65

70

cag ggc aaa gtg gtt caa gtt atc cgg cag cga aac tgg gtg gtc gtg 353
Gln Gly Lys Val Val Gln Val Ile Arg Gln Arg Asn Trp Val Val Val

75

80

85

gga ggg ctg aac aca cat tac cgc tac att ggc aag acc atg gat tac 401
Gly Gly Leu Asn Thr His Tyr Arg Tyr Ile Gly Lys Thr Met Asp Tyr

90

95

100

cgg gga acc atg atc cct agt gaa gcc ccc ttg ctc cac cgc cag gtc 449
Arg Gly Thr Met Ile Pro Ser Glu Ala Pro Leu Leu His Arg Gln Val

105

110

115

120

aaa ctt gtg gat cct atg gac agg aaa ccc act gag atc gag tgg aga 497

Lys Leu Val Asp Pro Met Asp Arg Lys Pro Thr Glu Ile Glu Trp Arg

125

130

135

ttt act gaa gca gga gag cgg gta cga gtc tcc aca cga tca ggg aga 545

Phe Thr Glu Ala Gly Glu Arg Val Arg Val Ser Thr Arg Ser Gly Arg

140

145

150

att atc cct aaa ccc gaa ttt ccc aga gct gat ggc atc gtc cct gaa 593

Ile Ile Pro Lys Pro Glu Phe Pro Arg Ala Asp Gly Ile Val Pro Glu

155

160

165

acg tgg att gat ggc ccc aaa gac aca tca gtg gaa gat gct tta gaa 641

Thr Trp Ile Asp Gly Pro Lys Asp Thr Ser Val Glu Asp Ala Leu Glu

170

175

180

aga acc tat gtg ccc tgt cta aag aca ctg cag gag gag gtg atg gag 689

Arg Thr Tyr Val Pro Cys Leu Lys Thr Leu Gln Glu Glu Val Met Glu

185

190

195

200

gcc atg ggg atc aag gag acc cgg aaa tac aag aag gtc tat tgg tat 737

Ala Met Gly Ile Lys Glu Thr Arg Lys Tyr Lys Lys Val Tyr Trp Tyr

205

210

215

tga gcctggggca gagcagctcc tccccaactt ctgtcccagc cttgaaggct 790

gaggcacttc tttttcagat gccataaag agcactttat gagtc 835

<210> 4

<211> 216

<212> PRT

<213> Homo sapiens

<400> 4

```

Met Arg Leu Ser Ala Leu Leu Ala Leu Ala Ser Lys Val Thr Leu Pro
  1           5           10           15
Pro His Tyr Arg Tyr Gly Met Ser Pro Pro Gly Ser Val Ala Asp Lys
      20           25           30
Arg Lys Asn Pro Pro Trp Ile Arg Arg Arg Pro Val Val Val Glu Pro
      35           40           45
Ile Ser Asp Glu Asp Trp Tyr Leu Phe Cys Gly Asp Thr Val Glu Ile
      50           55           60
Leu Glu Gly Lys Asp Ala Gly Lys Gln Gly Lys Val Val Gln Val Ile
      65           70           75           80
Arg Gln Arg Asn Trp Val Val Val Gly Gly Leu Asn Thr His Tyr Arg
      85           90           95
Tyr Ile Gly Lys Thr Met Asp Tyr Arg Gly Thr Met Ile Pro Ser Glu
      100          105          110
Ala Pro Leu Leu His Arg Gln Val Lys Leu Val Asp Pro Met Asp Arg
      115          120          125
Lys Pro Thr Glu Ile Glu Trp Arg Phe Thr Glu Ala Gly Glu Arg Val
      130          135          140
Arg Val Ser Thr Arg Ser Gly Arg Ile Ile Pro Lys Pro Glu Phe Pro
      145          150          155          160
Arg Ala Asp Gly Ile Val Pro Glu Thr Trp Ile Asp Gly Pro Lys Asp
      165          170          175

```

Thr Ser Val Glu Asp Ala Leu Glu Arg Thr Tyr Val Pro Cys Leu Lys
 180 185 190
 Thr Leu Gln Glu Glu Val Met Glu Ala Met Gly Ile Lys Glu Thr Arg
 195 200 205
 Lys Tyr Lys Lys Val Tyr Trp Tyr
 210 215

<210> 5

<211> 2465

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (71)..(1261)

<400> 5

aaactcgtca tttcctccag ctagaggagc tcaactgatc tgttttcttt cgcccagcca 60

aaatcacaga atg aag gcg gtg aag agc gaa cgg gag cga ggg agc cgg 109

Met Lys Ala Val Lys Ser Glu Arg Glu Arg Gly Ser Arg

1

5

10

cga aga cac cgg gac ggg gac gtg gtg ctg ccg gcg ggg gtg gtg gtg 157

Arg Arg His Arg Asp Gly Asp Val Val Leu Pro Ala Gly Val Val Val

15

20

25

aag cag gag cgt ctc agc cca gaa gtc gca cct ccc gcc cac cgc cgt 205
 Lys Gln Glu Arg Leu Ser Pro Glu Val Ala Pro Pro Ala His Arg Arg
 30 35 40 45

ccg gac cac tcc ggt ggt agc ccg tct ccg ccg acc agc gag ccg gcc 253
 Pro Asp His Ser Gly Gly Ser Pro Ser Pro Pro Thr Ser Glu Pro Ala
 50 55 60

cgc tcg ggc cac cgc ggg aac cga gcc cga gga gtt agc cgg tcc cca 301
 Arg Ser Gly His Arg Gly Asn Arg Ala Arg Gly Val Ser Arg Ser Pro
 65 70 75

ccc aaa aag aaa aac aag gcc tca ggg aga aga agc aag tct cct cgc 349
 Pro Lys Lys Lys Asn Lys Ala Ser Gly Arg Arg Ser Lys Ser Pro Arg
 80 85 90

agt aag aga aac cga agt cct cac cac tca aca gtc aaa gtg aag cag 397
 Ser Lys Arg Asn Arg Ser Pro His His Ser Thr Val Lys Val Lys Gln
 95 100 105

gag cgt gag gat cat ccc cgg aga gga cgg gag gat cgg cag cac agg 445
 Glu Arg Glu Asp His Pro Arg Arg Gly Arg Glu Asp Arg Gln His Arg
 110 115 120 125

gaa cca tca gaa cag gaa cac agg aga gct agg aac agt gac cgg gac 493
 Glu Pro Ser Glu Gln Glu His Arg Arg Ala Arg Asn Ser Asp Arg Asp
 130 135 140

aga cac cgg ggc cat tcc cac caa agg aga acg tct aac gag agg cct 541

Arg His Arg Gly His Ser His Gln Arg Arg Thr Ser Asn Glu Arg Pro
145 150 155

ggg agt ggg cag ggt cag gga cgg gat cga gac act cag aac ctg cag 589
Gly Ser Gly Gln Gly Gln Gly Arg Asp Arg Asp Thr Gln Asn Leu Gln
160 165 170

gct cag gaa gaa gag cgg gag ttt tat aat gcc agg cga cgg gag cat 637
Ala Gln Glu Glu Glu Arg Glu Phe Tyr Asn Ala Arg Arg Arg Glu His
175 180 185

cgc cag agg aat gac gtt ggt ggt ggc ggc agt gag tct cag gag ttg 685
Arg Gln Arg Asn Asp Val Gly Gly Gly Gly Ser Glu Ser Gln Glu Leu
190 195 200 205

gtt cct cgg cct ggt ggc aac aac aaa gaa aaa gag gtg ccc gct aaa 733
Val Pro Arg Pro Gly Gly Asn Asn Lys Glu Lys Glu Val Pro Ala Lys
210 215 220

gaa aaa cca agc ttt gaa ctt tct ggg gca ctt ctt gag gac acc aac 781
Glu Lys Pro Ser Phe Glu Leu Ser Gly Ala Leu Leu Glu Asp Thr Asn
225 230 235

act ttc cgg ggt gta gtc att aaa tat agt gag ccc cca gaa gca cgt 829
Thr Phe Arg Gly Val Val Ile Lys Tyr Ser Glu Pro Pro Glu Ala Arg
240 245 250

atc ccc aaa aaa cgg tgg cgt ctc tac cca ttt aaa aat gat gag gtg 877
Ile Pro Lys Lys Arg Trp Arg Leu Tyr Pro Phe Lys Asn Asp Glu Val

255	260	265	
ctt cca gtc atg tac ata cat cga cag agt gcg tac cta ctg ggt cga			925
Leu Pro Val Met Tyr Ile His Arg Gln Ser Ala Tyr Leu Leu Gly Arg			
270	275	280	285
cac cgc cgc att gca gac att cca att gat cac ccg tct tgt tca aag			973
His Arg Arg Ile Ala Asp Ile Pro Ile Asp His Pro Ser Cys Ser Lys			
	290	295	300
cag cat gcg gtc ttt caa tat cgg ctt gtg gaa tat acc cgt gct gat			1021
Gln His Ala Val Phe Gln Tyr Arg Leu Val Glu Tyr Thr Arg Ala Asp			
	305	310	315
ggc aca gtt ggc cga aga gtg aag ccc tac atc att gac ctt ggc tca			1069
Gly Thr Val Gly Arg Arg Val Lys Pro Tyr Ile Ile Asp Leu Gly Ser			
	320	325	330
ggc aat gga acc ttc tta aac aac aaa cgt att gag cca cag aga tac			1117
Gly Asn Gly Thr Phe Leu Asn Asn Lys Arg Ile Glu Pro Gln Arg Tyr			
	335	340	345
tat gaa cta aaa gaa aag gat gta ctc aaa ttt gga ttc agt agc aga			1165
Tyr Glu Leu Lys Glu Lys Asp Val Leu Lys Phe Gly Phe Ser Ser Arg			
350	355	360	365
gaa tac gtc ttg ctc cat gag tcg tcg gac act tct gaa ata gac agg			1213
Glu Tyr Val Leu Leu His Glu Ser Ser Asp Thr Ser Glu Ile Asp Arg			
	370	375	380

aaa gat gac gag gat gag gag gag gag gaa gaa gtg tct gac agc tag 1261

Lys Asp Asp Glu Asp Glu Glu Glu Glu Glu Val Ser Asp Ser

385

390

395

caaactaaga acccaaacta ttgatacag gtttccttct tggaagtctt tgattgactc 1321

agagagcact atggtggtgg gtccagcact atggtgctct ctgtaatgcc tcttactgcc 1381

ttaagtcttt cctctgttgc tgaccagatt gtgttaccat ttgaatacac tgactaatgt 1441

ttgttaaact ttttctgtgg caccttggcc acatgcctgc aggcatattgt tttcagaaca 1501

gtctcaccaa ttacaacaca ccgtgtttta gtagaagtgt tgttggtttta gttggtgctt 1561

tcagaactgc tgcctaggaa actataaacc cttggttaag gggaaatcat ggcttgttct 1621

ctttgtacag ttactttatt tatataggta ttaagctttg tggaccaggt gtttttcttt 1681

tggggcgaac ccctgagcag agaatcttac taggctttgg ttatcaccaa aacaacctcc 1741

agtatatacc aaagctttga cttgtttgag ctcttgagct tagaagttga ttttgactt 1801

atTTTTTTgg ggggtgggaa tgtactgcag tcagtaaaca ttattgactg ttttaactta 1861

acagatgctt tatggcacct gctcaagccc gtgactgtac agaaggatcc tggttgctac 1921

cagtgggtgc tgattcagca tcacaagtga ctgaaattgg ctgtggatct gttctttgtg 1981

aaagaattcc tgatttctcc atggagcatg tacacaacaa ttttgatcat attaactgta 2041
 cttcagtttt gcatttttat tcaaatgtta tctctttttt tctttgagaa ataaactgtc 2101
 actgatgtga cagcgttctt tctttattct aataacatgt atagatctaa agcaggttgt 2161
 gttgtttaca tgtttctaca catttcatcc tttaaaaagt tgttgagaga ggttgtattt 2221
 accttcccaa ggttggaag caggggaatt tcccagtgtc ctagttttcc accagaggaa 2281
 tatgtgtaag tagcaaagta ttgctgctt acatatagtg tgtatgtatg tatatatgta 2341
 aattgtgtgt taaagagctg atactgattt tcatatgaca atgttaggca aaggcctccc 2401
 tgcattigaa gagcaggttt tcatttatat gtatttttgg gataaaaaaa taaatttgta 2461
 atat 2465

<210> 6

<211> 396

<212> PRT

<213> Homo sapiens

<400> 6

Met Lys Ala Val Lys Ser Glu Arg Glu Arg Gly Ser Arg Arg Arg His

1

5

10

15

Arg Asp Gly Asp Val Val Leu Pro Ala Gly Val Val Val Lys Gln Glu

20

25

30

Arg Leu Ser Pro Glu Val Ala Pro Pro Ala His Arg Arg Pro Asp His
 35 40 45
 Ser Gly Gly Ser Pro Ser Pro Pro Thr Ser Glu Pro Ala Arg Ser Gly
 50 55 60
 His Arg Gly Asn Arg Ala Arg Gly Val Ser Arg Ser Pro Pro Lys Lys
 65 70 75 80
 Lys Asn Lys Ala Ser Gly Arg Arg Ser Lys Ser Pro Arg Ser Lys Arg
 85 90 95
 Asn Arg Ser Pro His His Ser Thr Val Lys Val Lys Gln Glu Arg Glu
 100 105 110
 Asp His Pro Arg Arg Gly Arg Glu Asp Arg Gln His Arg Glu Pro Ser
 115 120 125
 Glu Gln Glu His Arg Arg Ala Arg Asn Ser Asp Arg Asp Arg His Arg
 130 135 140
 Gly His Ser His Gln Arg Arg Thr Ser Asn Glu Arg Pro Gly Ser Gly
 145 150 155 160
 Gln Gly Gln Gly Arg Asp Arg Asp Thr Gln Asn Leu Gln Ala Gln Glu
 165 170 175
 Glu Glu Arg Glu Phe Tyr Asn Ala Arg Arg Arg Glu His Arg Gln Arg
 180 185 190
 Asn Asp Val Gly Gly Gly Gly Ser Glu Ser Gln Glu Leu Val Pro Arg
 195 200 205
 Pro Gly Gly Asn Asn Lys Glu Lys Glu Val Pro Ala Lys Glu Lys Pro
 210 215 220
 Ser Phe Glu Leu Ser Gly Ala Leu Leu Glu Asp Thr Asn Thr Phe Arg
 225 230 235 240
 Gly Val Val Ile Lys Tyr Ser Glu Pro Pro Glu Ala Arg Ile Pro Lys
 245 250 255
 Lys Arg Trp Arg Leu Tyr Pro Phe Lys Asn Asp Glu Val Leu Pro Val

260	265	270
Met Tyr Ile His Arg Gln Ser Ala Tyr Leu Leu Gly Arg His Arg Arg		
275	280	285
Ile Ala Asp Ile Pro Ile Asp His Pro Ser Cys Ser Lys Gln His Ala		
290	295	300
Val Phe Gln Tyr Arg Leu Val Glu Tyr Thr Arg Ala Asp Gly Thr Val		
305	310	315
Gly Arg Arg Val Lys Pro Tyr Ile Ile Asp Leu Gly Ser Gly Asn Gly		
325	330	335
Thr Phe Leu Asn Asn Lys Arg Ile Glu Pro Gln Arg Tyr Tyr Glu Leu		
340	345	350
Lys Glu Lys Asp Val Leu Lys Phe Gly Phe Ser Ser Arg Glu Tyr Val		
355	360	365
Leu Leu His Glu Ser Ser Asp Thr Ser Glu Ile Asp Arg Lys Asp Asp		
370	375	380
Glu Asp Glu Glu Glu Glu Glu Glu Val Ser Asp Ser		
385	390	395

<210> 7

<211> 3600

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (149)..(1504)

<400> 7

aatccagcgc agccgggaga cagatgcgag gcggcgggtca gcaggtgccg aaccacaggc 60

caggcttccg tggccagcag ccctagagga atggccatcc tgtccctgcg agcccctggg 120

ccctggcagg cgatgcaggt gggatgct atg gaa tat gat gag aag ctg gcc 172

Met Glu Tyr Asp Glu Lys Leu Ala

1

5

cgt ttc cgg cag gcc cac ctc aac ccc ttc aac aag cag tct ggg ccg 220

Arg Phe Arg Gln Ala His Leu Asn Pro Phe Asn Lys Gln Ser Gly Pro

10

15

20

aga cag cat gag cag ggc cct ggg gag gag gtc ccg gac gtc act cct 268

Arg Gln His Glu Gln Gly Pro Gly Glu Glu Val Pro Asp Val Thr Pro

25

30

35

40

gaa gag gcc ctg cct gag ctg ccc cct ggg gag ccg gaa ttc cgc tgc 316

Glu Glu Ala Leu Pro Glu Leu Pro Pro Gly Glu Pro Glu Phe Arg Cys

45

50

55

cct gaa cgc gtg atg gat ctc ggc ctg tct gag gac cac ttc tcc cgc 364

Pro Glu Arg Val Met Asp Leu Gly Leu Ser Glu Asp His Phe Ser Arg

60

65

70

cct gtg ggt ctg ttc ctg gcc tct gac gtc cag cag ctg cgg cag gcg 412

Pro Val Gly Leu Phe Leu Ala Ser Asp Val Gln Gln Leu Arg Gln Ala

75

80

85

atc gag gag tgc aag cag gtg att ctg gag ctg ccc gag cag tcg gag 460
 Ile Glu Glu Cys Lys Gln Val Ile Leu Glu Leu Pro Glu Gln Ser Glu
 90 95 100

aag cag aag gat gcc gtg gtg cga ctc atc cac ctc cgg ctg aag ctc 508
 Lys Gln Lys Asp Ala Val Val Arg Leu Ile His Leu Arg Leu Lys Leu
 105 110 115 120

cag gag ctg aag gac ccc aat gag gat gag cca aac atc cga gtg ctc 556
 Gln Glu Leu Lys Asp Pro Asn Glu Asp Glu Pro Asn Ile Arg Val Leu
 125 130 135

ctt gag cac cgc ttt tac aag gag aag agc aag agc gtc aag cag acc 604
 Leu Glu His Arg Phe Tyr Lys Glu Lys Ser Lys Ser Val Lys Gln Thr
 140 145 150

tgt gac aag tgt aac acc atc atc tgg ggg ctc att cag acc tgg tac 652
 Cys Asp Lys Cys Asn Thr Ile Ile Trp Gly Leu Ile Gln Thr Trp Tyr
 155 160 165

acc tgc aca ggg tgt tat tac cgc tgt cac agt aag tgc ttg aac ctc 700
 Thr Cys Thr Gly Cys Tyr Tyr Arg Cys His Ser Lys Cys Leu Asn Leu
 170 175 180

atc tcc aag ccc tgt gtg agc tcc aaa gtc agc cac caa gct gaa tac 748
 Ile Ser Lys Pro Cys Val Ser Ser Lys Val Ser His Gln Ala Glu Tyr
 185 190 195 200

gaa ctg aac atc tgc cct gag aca ggg ctg gac agc cag gat tac cgc 796

Glu Leu Asn Ile Cys Pro Glu Thr Gly Leu Asp Ser Gln Asp Tyr Arg
205 210 215

tgt gcc gag tgc cgg gcg ccc atc tct ctg cgg ggt gtg ccc agt gag 844
Cys Ala Glu Cys Arg Ala Pro Ile Ser Leu Arg Gly Val Pro Ser Glu
220 225 230

gcc agg cag tgc gac tac acc ggc cag tac tac tgc agc cac tgc cac 892
Ala Arg Gln Cys Asp Tyr Thr Gly Gln Tyr Tyr Cys Ser His Cys His
235 240 245

tgg aac gac ctg gct gtg atc cct gca cgc gtt gta cac aac tgg gac 940
Trp Asn Asp Leu Ala Val Ile Pro Ala Arg Val Val His Asn Trp Asp
250 255 260

ttt gag cct cga aag gtt tct cgc tgc agc atg cgc tac ctg gcg ctg 988
Phe Glu Pro Arg Lys Val Ser Arg Cys Ser Met Arg Tyr Leu Ala Leu
265 270 275 280

atg gtg tct cgg ccc gta ctc agg ctc cgg gag atc aac cct ctg ctg 1036
Met Val Ser Arg Pro Val Leu Arg Leu Arg Glu Ile Asn Pro Leu Leu
285 290 295

ttc agc tac gtg gag gag ctg gtg gag att cgc aag ctg cgc cag gac 1084
Phe Ser Tyr Val Glu Glu Leu Val Glu Ile Arg Lys Leu Arg Gln Asp
300 305 310

atc ctg ctc atg aag ccg tac ttc atc acc tgc agg gag gcc atg gag 1132
Ile Leu Leu Met Lys Pro Tyr Phe Ile Thr Cys Arg Glu Ala Met Glu

315	320	325	
gct cgt ctg ctg ctg cag ctc cag gat cgg cag cat ttt gtg gag aac			1180
Ala Arg Leu Leu Leu Gln Leu Gln Asp Arg Gln His Phe Val Glu Asn			
330	335	340	
gac gag atg tac tct gtc cag gac ctc ctg gac gtg cat gcc ggc cgc			1228
Asp Glu Met Tyr Ser Val Gln Asp Leu Leu Asp Val His Ala Gly Arg			
345	350	355	360
ctg ggc tgc tcg ctc acc gag atc cac acg ctc ttc gcc aag cac atc			1276
Leu Gly Cys Ser Leu Thr Glu Ile His Thr Leu Phe Ala Lys His Ile			
365	370	375	
aag ctg gac tgc gag cgg tgc cag gcc aag ggc ttc gtg tgt gag ctc			1324
Lys Leu Asp Cys Glu Arg Cys Gln Ala Lys Gly Phe Val Cys Glu Leu			
380	385	390	
tgc aga gag ggc gac gtg ctg ttc ccg ttc gac agc cac acg tct gtg			1372
Cys Arg Glu Gly Asp Val Leu Phe Pro Phe Asp Ser His Thr Ser Val			
395	400	405	
tgc gcc gac tgc tcc gcg gtc ttc cac agg gac tgc tac tac gac aac			1420
Cys Ala Asp Cys Ser Ala Val Phe His Arg Asp Cys Tyr Tyr Asp Asn			
410	415	420	
tcc acc act tgt ccc aag tgt gcc cgg ctc agc ctg agg aag cag tcg			1468
Ser Thr Thr Cys Pro Lys Cys Ala Arg Leu Ser Leu Arg Lys Gln Ser			
425	430	435	440

ctc ttc cag gag cca ggt ccc gat gtg gag gcc tag cgccgaggaa 1514

Leu Phe Gln Glu Pro Gly Pro Asp Val Glu Ala

445

450

cagtgtctggg caccctgcct ggcccgccag gaccacacct gccaacatca agttgttcct 1574

tctgtctccgg agacccttgg ggtgcggccc tggccccctc caccctgtct gggccagagc 1634

gggtgggcag tgtcaaggcc cgctgtctcc caggtgttgc ctgggactcg gggcggctgc 1694

acctggctgt cacctgggtg tctgtctgtg aggggtcctt gcgtggcccc catccttccc 1754

ccaatgcaga actccatggg caggagctg gggggacatc tcacctcccc catggcacag 1814

agccctccac acccctggac caggcatcc gggccctaga aattccacag ctcccgctct 1874

ggccaccctg gaagctcatc aggccaagac ccggacagag cticagagga gtgttgagtg 1934

acacctgagg atgcggctgc acacactcag ccaagggccg agtctcacct gcggtgggggt 1994

ttcggctctg cctgggggct ccatcccttt cagccactcg tggccttggg gatttcttgt 2054

tgtccccagc tgggactgtt cacagttgtc acctgcagac ctgcctctcc ctggcctgag 2114

gttcaaaggc ctcatcgat ggtcagtaca gtggggtcac ctgttgtttc tatacaacag 2174

cagggaaggg gccatggagc ttttccctgc tgggtgtctc tgctttggcc cagccacacct 2234

ttccctggtgc tccaagctag gaggctgtgg cccagcctg aggagggtgt cctggcctcc 2294
 aggtgtgcag caggggctgt gtgctggggg aggttccagt taggcgatgg gacccctgcag 2354
 tggctctggtg gcatttcttg gaaccagatt tacctgagga gctctgtcct gctccctgtg 2414
 gagggctcca gatagctcag aaatgaccag ccaatggcct ttgttttggg ggcctgaggt 2474
 caagagagct gagagtattc gctcgactga gcacattcag gaagatcagg gcaggcgtgt 2534
 gggaggtccc tcactccacg ggacagaggc ccttgacag cagaggaaac ctacagctct 2594
 gggtaggggg acacttggct ttggtgtttg cactttacag atcctgcggt ccacgagggg 2654
 cctcaggaga ggacgtgtca ggacgtggct tcccagcctt ctgccttggg cagtgggggt 2714
 gctcctgtct gtccttttcc cccacaccct ggactgtgct tggctgttgg tgcacatggt 2774
 tggcacacgg tgggcagagg gcagagaatg ccactgcttg gttattggtc ccctttgacc 2834
 aggaaacca agaggagaca cctcagtcag cagaaaggcc acctggctca ctggctcatt 2894
 ccaggagtgg gagagacggc agggctctct ctttgtcttc cggcatcagg aaggggatgg 2954
 tgtccactcc ccactgtggt ggcttttaggc aaggttctta ttgtctgctc tgcctcggtt 3014
 tccccatctg gaaaatgggg gcaggggtcc tgacctacct caggtggaac ggtgagcagg 3074
 gaacatgtcg gagtccctca gagaatgtga tgtgaggttg gatcaacagt gtgggttcct 3134

gtcctgtttc cccttcctct ttggggctga ggaggaggtt aaaggccaaa tgctgtttcc 3194
 caacacccca aagtctgcac acgtctcatg aatgcatcac atttctgtca tatggatatt 3254
 agccattccg aaatctgtgt aatcaacttc acattattca agttacaaat cactgtgtcc 3314
 atagaaaaac tgtgctggta ttgctggac aaagggttgg gccccttta tttttacctg 3374
 ccaccagca tctccccac ctgccccttc tgggtgacac agccggtaaa cggaatcacg 3434
 tatggttctt tctgtgggtc tgtggcacag caggaagagc ccggtgccgc cagcaccttg 3494
 tggaagacca cacatgggtg gtcccacagc atgggaccag gctggcctga gggatgcccc 3554
 gttgtaacaa tgctgctgtc actgtctcat taaatataca tccttt 3600

<210> 8

<211> 451

<212> PRT

<213> Homo sapiens

<400> 8

Met	Glu	Tyr	Asp	Glu	Lys	Leu	Ala	Arg	Phe	Arg	Gln	Ala	His	Leu	Asn
1				5				10						15	
Pro	Phe	Asn	Lys	Gln	Ser	Gly	Pro	Arg	Gln	His	Glu	Gln	Gly	Pr	Gly
				20				25						30	
Glu	Glu	Val	Pro	Asp	Val	Thr	Pro	Glu	Glu	Ala	Leu	Pro	Glu	Leu	Pro

35	40	45
Pro Gly Glu Pro Glu Phe Arg Cys Pro Glu Arg Val Met Asp Leu Gly		
50	55	60
Leu Ser Glu Asp His Phe Ser Arg Pro Val Gly Leu Phe Leu Ala Ser		
65	70	75
Asp Val Gln Gln Leu Arg Gln Ala Ile Glu Glu Cys Lys Gln Val Ile		
85	90	95
Leu Glu Leu Pro Glu Gln Ser Glu Lys Gln Lys Asp Ala Val Val Arg		
100	105	110
Leu Ile His Leu Arg Leu Lys Leu Gln Glu Leu Lys Asp Pro Asn Glu		
115	120	125
Asp Glu Pro Asn Ile Arg Val Leu Leu Glu His Arg Phe Tyr Lys Glu		
130	135	140
Lys Ser Lys Ser Val Lys Gln Thr Cys Asp Lys Cys Asn Thr Ile Ile		
145	150	155
Trp Gly Leu Ile Gln Thr Trp Tyr Thr Cys Thr Gly Cys Tyr Tyr Arg		
165	170	175
Cys His Ser Lys Cys Leu Asn Leu Ile Ser Lys Pro Cys Val Ser Ser		
180	185	190
Lys Val Ser His Gln Ala Glu Tyr Glu Leu Asn Ile Cys Pro Glu Thr		
195	200	205
Gly Leu Asp Ser Gln Asp Tyr Arg Cys Ala Glu Cys Arg Ala Pro Ile		
210	215	220
Ser Leu Arg Gly Val Pro Ser Glu Ala Arg Gln Cys Asp Tyr Thr Gly		
225	230	235
Gln Tyr Tyr Cys Ser His Cys His Trp Asn Asp Leu Ala Val Ile Pro		
245	250	255
Ala Arg Val Val His Asn Trp Asp Phe Glu Pro Arg Lys Val Ser Arg		
260	265	270

Cys Ser Met Arg Tyr Leu Ala Leu Met Val Ser Arg Pro Val Leu Arg			
275	280	285	
Leu Arg Glu Ile Asn Pro Leu Leu Phe Ser Tyr Val Glu Glu Leu Val			
290	295	300	
Glu Ile Arg Lys Leu Arg Gln Asp Ile Leu Leu Met Lys Pro Tyr Phe			
305	310	315	320
Ile Thr Cys Arg Glu Ala Met Glu Ala Arg Leu Leu Leu Gln Leu Gln			
325	330	335	
Asp Arg Gln His Phe Val Glu Asn Asp Glu Met Tyr Ser Val Gln Asp			
340	345	350	
Leu Leu Asp Val His Ala Gly Arg Leu Gly Cys Ser Leu Thr Glu Ile			
355	360	365	
His Thr Leu Phe Ala Lys His Ile Lys Leu Asp Cys Glu Arg Cys Gln			
370	375	380	
Ala Lys Gly Phe Val Cys Glu Leu Cys Arg Glu Gly Asp Val Leu Phe			
385	390	395	400
Pro Phe Asp Ser His Thr Ser Val Cys Ala Asp Cys Ser Ala Val Phe			
405	410	415	
His Arg Asp Cys Tyr Tyr Asp Asn Ser Thr Thr Cys Pro Lys Cys Ala			
420	425	430	
Arg Leu Ser Leu Arg Lys Gln Ser Leu Phe Gln Glu Pro Gly Pro Asp			
435	440	445	
Val Glu Ala			
450			

<210> 9

<211> 442

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (12)..(353)

<400> 9

gttctacagc t atg gcc ggg cca gct gca gct ttc cgc cgc ttg ggc gcc 50

Met Ala Gly Pro Ala Ala Ala Phe Arg Arg Leu Gly Ala

1

5

10

ttg tcc gga gct gcg gcc tta ggc ttc gct tcc tac ggg gcg cac ggc 98

Leu Ser Gly Ala Ala Ala Leu Gly Phe Ala Ser Tyr Gly Ala His Gly

15

20

25

gcc caa ttc cca gat gcc tac ggg aag gag ctg ttt gac aag gcc aac 146

Ala Gln Phe Pro Asp Ala Tyr Gly Lys Glu Leu Phe Asp Lys Ala Asn

30

35

40

45

aaa cac cac ttc tta cac agc ctg gcc ctg tta ggg gtg ccc cat tgc 194

Lys His His Phe Leu His Ser Leu Ala Leu Leu Gly Val Pro His Cys

50

55

60

aga aag cca ctc tgg gct ggg tta ttg cta gct tcc gga acg acc tta 242

Arg Lys Pro Leu Trp Ala Gly Leu Leu Leu Ala Ser Gly Thr Thr Leu

65

70

75

ttc tgc acc agc ttt tac tac cag gct ctg agt gga gac ccc agc atc 290

Phe Cys Thr Ser Phe Tyr Tyr Gln Ala Leu Ser Gly Asp Pro Ser Ile

80

85

90

cag act ttg gcc cct gcg gga ggg acc ctg cta ctc ttg ggc tgg ctt 338

Gln Thr Leu Ala Pro Ala Gly Gly Thr Leu Leu Leu Leu Gly Trp Leu

95

100

105

gcc ttg gct ctt tga gctccctttt gcttaattac tgggttttct gggcagtttt 393

Ala Leu Ala Leu

110

tttttttaaa gagttggagt aagaagagga ttaaaaagga aaggcaaat 442

<210> 10

<211> 113

<212> PRT

<213> Homo sapiens

<400> 10

Met Ala Gly Pro Ala Ala Ala Phe Arg Arg Leu Gly Ala Leu Ser Gly

1

5

10

15

Ala Ala Ala Leu Gly Phe Ala Ser Tyr Gly Ala His Gly Ala Gln Phe

20

25

30

Pro Asp Ala Tyr Gly Lys Glu Leu Phe Asp Lys Ala Asn Lys His His

35

40

45

Phe Leu His Ser Leu Ala Leu Leu Gly Val Pro His Cys Arg Lys Pro

50

55

60

Leu Trp Ala Gly Leu Leu Leu Ala Ser Gly Thr Thr Leu Phe Cys Thr

65	70	75	80
Ser Phe Tyr Tyr Gln Ala Leu Ser Gly Asp Pro Ser Ile Gln Thr Leu			
	85	90	95
Ala Pro Ala Gly Gly Thr Leu Leu Leu Leu Gly Trp Leu Ala Leu Ala			
	100	105	110
Leu			

<210> 11

<211> 726

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (12)..(680)

<400> 11

gcacagccaa g atg gcg gcg tcc gtg cga cag gca cgc agc cta cta ggt 50

Met Ala Ala Ser Val Arg Gln Ala Arg Ser Leu Leu Gly

1

5

10

gtg gcg gcg acc ctg gcc ccg ggt tcc cgt ggc tac cgg gcg cgg ccg 98

Val Ala Ala Thr Leu Ala Pro Gly Ser Arg Gly Tyr Arg Ala Arg Pro

15

20

25

ccc ccg cgc cgc agg ccg gga ccc cgg tgg cca gac ccc gag gac ctc 146

Pro Pro Arg Arg Arg Pro Gly Pro Arg Trp Pro Asp Pro Glu Asp Leu

30	35	40	45	
ctg acc ccg cgg tgg cag ctg gga ccg cgc tac gcg gct aag cag ttc	194			
Leu Thr Pro Arg Trp Gln Leu Gly Pro Arg Tyr Ala Ala Lys Gln Phe				
50	55	60		
gcg cgt tac ggc gcc gcc tcc ggg gtg gtc ccc ggt tcg tta tgg ccg	242			
Ala Arg Tyr Gly Ala Ala Ser Gly Val Val Pro Gly Ser Leu Trp Pro				
65	70	75		
tcg ccg gag cag ctg cgg gag ctg gag gcc gaa gaa cgc gaa tgg tac	290			
Ser Pro Glu Gln Leu Arg Glu Leu Glu Ala Glu Glu Arg Glu Trp Tyr				
80	85	90		
ccg agc ctg gcg acc atg cag gag tcg ctg cgg gtg aag cag ctg gcc	338			
Pro Ser Leu Ala Thr Met Gln Glu Ser Leu Arg Val Lys Gln Leu Ala				
95	100	105		
gaa gag cag aag cgt cgg gag agg gag cag cac atc gca gag tgc atg	386			
Glu Glu Gln Lys Arg Arg Glu Arg Glu Gln His Ile Ala Glu Cys Met				
110	115	120	125	
gcc aag atg cca cag atg att gtg aac tgg cag cag cag cag cgg gag	434			
Ala Lys Met Pro Gln Met Ile Val Asn Trp Gln Gln Gln Gln Arg Glu				
130	135	140		
aac tgg gag aag gcc cag gct gac aag gag agg agg gcc cga ctg cag	482			
Asn Trp Glu Lys Ala Gln Ala Asp Lys Glu Arg Arg Ala Arg Leu Gln				
145	150	155		

gct gag gcc cag gag ctc ctg ggc tac cag gtg gac cca agg agt gcc 530

Ala Glu Ala Gln Glu Leu Leu Gly Tyr Gln Val Asp Pro Arg Ser Ala

160

165

170

cgc ttc cag gag ctg ctc cag gac cta gag aag aag gag cgc aag cgc 578

Arg Phe Gln Glu Leu Leu Gln Asp Leu Glu Lys Lys Glu Arg Lys Arg

175

180

185

ctc aag gag gaa aaa cag aaa cgg aag aag gag gcg cga gct gct gca 626

Leu Lys Glu Glu Lys Gln Lys Arg Lys Lys Glu Ala Arg Ala Ala Ala

190

195

200

205

ttg gct gca gct gtg gct caa gac cca gca gcc tct ggg gca ccc agc 674

Leu Ala Ala Ala Val Ala Gln Asp Pro Ala Ala Ser Gly Ala Pro Ser

210

215

220

tcc tga ggctttgtcc ctcccaata aagcctgcta cctggcagta cccctg 726

Ser

<210> 12

<211> 222

<212> PRT

<213> Homo sapiens

<400> 12

Met Ala Ala Ser Val Arg Gln Ala Arg Ser Leu Leu Gly Val Ala Ala

1

5

10

15

5 1

<210> 13

<211> 1120

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (208)..(1044)

<400> 13

ccttttcttc cgcacggttg gaggaggtcg gctggttatac gggagttgga gggctgaggt 60

cgggagggtg gtgtgtacag agctctagga ctacgcacc aggccagtcg cgggttttgg 120

gccgaggcct gggttacaag cagcaagtgc gcggttgggg ccactgcgag gccgttttag 180

aaaactgttt aaaacaaaga gcaattg atg gat aaa tca gga ata gat tct ctt 234

Met Asp Lys Ser Gly Ile Asp Ser Leu

1

5

gac cat gtg aca tct gat gct gtg gaa ctt gca aat cga agt gat aac 282

Asp His Val Thr Ser Asp Ala Val Glu Leu Ala Asn Arg Ser Asp Asn

10

15

20

25

tct tct gat agc agc tta ttt aaa act cag tgt atc cct tac tca cct 330

Ser Ser Asp Ser Ser Leu Phe Lys Thr Gln Cys Ile Pro Tyr Ser Pro

30

35

40

aaa ggg gag aaa aga aac ccc att cga aaa ttt gtt cgt aca cct gaa 378

Lys Gly Glu Lys Arg Asn Pro Ile Arg Lys Phe Val Arg Thr Pro Glu

45

50

55

agt gtt cac gca agt gat tca tca agt gac tca tct ttt gaa cca ata 426

Ser Val His Ala Ser Asp Ser Ser Ser Asp Ser Ser Phe Glu Pro Ile

60

65

70

cca ttg act ata aaa gct att ttt gaa aga ttc aag aac agg aaa aag 474

Pro Leu Thr Ile Lys Ala Ile Phe Glu Arg Phe Lys Asn Arg Lys Lys

75

80

85

aga tat aaa aaa aag aaa aag agg agg tac cag cca aca gga aga cca 522

Arg Tyr Lys Lys Lys Lys Lys Arg Arg Tyr Gln Pro Thr Gly Arg Pro

90

95

100

105

cgg gga aga cca gaa gga agg aga aat cct ata tac tca cta ata gat 570

Arg Gly Arg Pro Glu Gly Arg Arg Asn Pro Ile Tyr Ser Leu Ile Asp

110

115

120

aag aag aaa caa ttt aga agc aga gga tct ggc ttc cca ttt tta gaa 618

Lys Lys Lys Gln Phe Arg Ser Arg Gly Ser Gly Phe Pro Phe Leu Glu

125

130

135

tca gag aat gaa aaa aac gca cct tgg aga aaa att tta acg ttt gag 666

Ser Glu Asn Glu Lys Asn Ala Pro Trp Arg Lys Ile Leu Thr Phe Glu

140

145

150

caa gct gtt gca aga gga ttt ttt aac tat att gaa aaa ctg aag tat 714

Gln Ala Val Ala Arg Gly Phe Phe Asn Tyr Ile Glu Lys Leu Lys Tyr

155	160	165	
gaa cac cac ctg aaa gaa tca ttg aag caa atg aat gtt ggt gaa gat			762
Glu His His Leu Lys Glu Ser Leu Lys Gln Met Asn Val Gly Glu Asp			
170	175	180	185
tta gaa aat gaa gat ttt gac agt cgt aga tac aaa ttt ttg gat gat			810
Leu Glu Asn Glu Asp Phe Asp Ser Arg Arg Tyr Lys Phe Leu Asp Asp			
	190	195	200
gat gga tcc att tct cct att gag gag tca aca gca gag gat gag gat			858
Asp Gly Ser Ile Ser Pro Ile Glu Glu Ser Thr Ala Glu Asp Glu Asp			
	205	210	215
gca aca cat ctt gaa gat aac gaa tgt gat atc aaa ttg gca ggg gat			906
Ala Thr His Leu Glu Asp Asn Glu Cys Asp Ile Lys Leu Ala Gly Asp			
	220	225	230
agt ttc ata gta agt tct gaa ttc cct gta aga ctg agt gta tac tta			954
Ser Phe Ile Val Ser Ser Glu Phe Pro Val Arg Leu Ser Val Tyr Leu			
	235	240	245
gaa gaa gag gat att act gaa gaa gct gct ttg tct aaa aag aga gct			1002
Glu Glu Glu Asp Ile Thr Glu Glu Ala Ala Leu Ser Lys Lys Arg Ala			
250	255	260	265
aca aaa gcc aaa aat act gga cag aga ggc ctg aaa atg tga			1044
Thr Lys Ala Lys Asn Thr Gly Gln Arg Gly Leu Lys Met			
	270	275	

caggatcatg aatgtcaaag gtgaagcata tagaaaaaac gacttcatag aaatgaataa 1104

agataaatgt ggatat 1120

<210> 14

<211> 278

<212> PRT

<213> Homo sapiens

<400> 14

Met	Asp	Lys	Ser	Gly	Ile	Asp	Ser	Leu	Asp	His	Val	Thr	Ser	Asp	Ala
1				5					10					15	
Val	Glu	Leu	Ala	Asn	Arg	Ser	Asp	Asn	Ser	Ser	Asp	Ser	Ser	Leu	Phe
			20					25						30	
Lys	Thr	Gln	Cys	Ile	Pro	Tyr	Ser	Pro	Lys	Gly	Glu	Lys	Arg	Asn	Pro
			35					40						45	
Ile	Arg	Lys	Phe	Val	Arg	Thr	Pro	Glu	Ser	Val	His	Ala	Ser	Asp	Ser
		50					55					60			
Ser	Ser	Asp	Ser	Ser	Phe	Glu	Pro	Ile	Pro	Leu	Thr	Ile	Lys	Ala	Ile
		65				70					75				80
Phe	Glu	Arg	Phe	Lys	Asn	Arg	Lys	Lys	Arg	Tyr	Lys	Lys	Lys	Lys	Lys
				85						90					95
Arg	Arg	Tyr	Gln	Pro	Thr	Gly	Arg	Pro	Arg	Gly	Arg	Pro	Glu	Gly	Arg
				100						105					110
Arg	Asn	Pro	Ile	Tyr	Ser	Leu	Ile	Asp	Lys	Lys	Lys	Gln	Phe	Arg	Ser
				115						120					125
Arg	Gly	Ser	Gly	Phe	Pro	Phe	Leu	Glu	Ser	Glu	Asn	Glu	Lys	Asn	Ala

130	135	140	
Pro Trp Arg Lys Ile Leu Thr Phe Glu Gln Ala Val Ala Arg Gly Phe			
145	150	155	160
Phe Asn Tyr Ile Glu Lys Leu Lys Tyr Glu His His Leu Lys Glu Ser			
	165	170	175
Leu Lys Gln Met Asn Val Gly Glu Asp Leu Glu Asn Glu Asp Phe Asp			
	180	185	190
Ser Arg Arg Tyr Lys Phe Leu Asp Asp Asp Gly Ser Ile Ser Pro Ile			
	195	200	205
Glu Glu Ser Thr Ala Glu Asp Glu Asp Ala Thr His Leu Glu Asp Asn			
	210	215	220
Glu Cys Asp Ile Lys Leu Ala Gly Asp Ser Phe Ile Val Ser Ser Glu			
225	230	235	240
Phe Pro Val Arg Leu Ser Val Tyr Leu Glu Glu Glu Asp Ile Thr Glu			
	245	250	255
Glu Ala Ala Leu Ser Lys Lys Arg Ala Thr Lys Ala Lys Asn Thr Gly			
	260	265	270
Gln Arg Gly Leu Lys Met			
275			

<210> 15

<211> 747

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (27)..(692)

<400> 15

ctctctcatg accccgctcc gggatt atg gcc ggg act ggg ctg ctg gcg ctg 53

Met Ala Gly Thr Gly Leu Leu Ala Leu

1

5

cgg acg ctg cca ggg ccc agc tgg gtg cga ggc tcg ggc cct tcc gtg 101

Arg Thr Leu Pro Gly Pro Ser Trp Val Arg Gly Ser Gly Pro Ser Val

10

15

20

25

ctg agc cgc ctg cag gac gcg gcc gtg gtg cgg cct ggc ttc ctg agc 149

Leu Ser Arg Leu Gln Asp Ala Ala Val Val Arg Pro Gly Phe Leu Ser

30

35

40

acg gca gag gag gag acg ctg agc cga gaa ctg gag ccc gag ctg cgc 197

Thr Ala Glu Glu Glu Thr Leu Ser Arg Glu Leu Glu Pro Glu Leu Arg

45

50

55

cgc cgc cgc tac gaa tac gat cac tgg gac gcg gcc atc cac ggc ttc 245

Arg Arg Arg Tyr Glu Tyr Asp His Trp Asp Ala Ala Ile His Gly Phe

60

65

70

cga gag aca gag aag tcg cgc tgg tca gaa gcc agc cgg gcc atc ctg 293

Arg Glu Thr Glu Lys Ser Arg Trp Ser Glu Ala Ser Arg Ala Ile Leu

75

80

85

cag cgc gtg cag gcg gcc gcc ttt ggc ccc ggc cag acc ctg ctc tcc 341

Gln Arg Val Gln Ala Ala Ala Phe Gly Pro Gly Gln Thr Leu Leu Ser

90	95	100	105	
tcc gtg cac gtg ctg gac ctg gaa gcc cgc ggc tac atc aag ccc cac	389			
Ser Val His Val Leu Asp Leu Glu Ala Arg Gly Tyr Ile Lys Pro His				
110	115	120		
gtg gac agc atc aag ttc tgc ggg gcc acc atc gcc ggc ctg tct ctc	437			
Val Asp Ser Ile Lys Phe Cys Gly Ala Thr Ile Ala Gly Leu Ser Leu				
125	130	135		
ctg tct ccc agc gtt atg cgg ctg gtg cac acc cag gag ccg ggg gag	485			
Leu Ser Pro Ser Val Met Arg Leu Val His Thr Gln Glu Pro Gly Glu				
140	145	150		
tgg ctg gaa ctc ttg ctg gag ccg ggc tcc ctc tac atc ctt agg ggc	533			
Trp Leu Glu Leu Leu Leu Glu Pro Gly Ser Leu Tyr Ile Leu Arg Gly				
155	160	165		
tca gcc cgt tat gac ttc tcc cat gag atc ctt cgg gat gaa gag tcc	581			
Ser Ala Arg Tyr Asp Phe Ser His Glu Ile Leu Arg Asp Glu Glu Ser				
170	175	180	185	
ttc ttt ggg gaa cgc cgg att ccc cgg ggc cgg cgc atc tcc gtg atc	629			
Phe Phe Gly Glu Arg Arg Ile Pro Arg Gly Arg Arg Ile Ser Val Ile				
190	195	200		
tgc cgc tcc ctc cct gag ggc atg ggg cca ggg gag tct gga cag ccg	677			
Cys Arg Ser Leu Pro Glu Gly Met Gly Pro Gly Glu Ser Gly Gln Pro				
205	210	215		

ccc cca gcc tgc tga cccccagctt tctacagaca ccagatttgt gaataaagtt 732

Pro Pro Ala Cys

220

ggggaatgga cagcc

747

<210> 16

<211> 221

<212> PRT

<213> Homo sapiens

<400> 16

Met Ala Gly Thr Gly Leu Leu Ala Leu Arg Thr Leu Pro Gly Pro Ser

1 5 10 15

Trp Val Arg Gly Ser Gly Pro Ser Val Leu Ser Arg Leu Gln Asp Ala

20 25 30

Ala Val Val Arg Pro Gly Phe Leu Ser Thr Ala Glu Glu Glu Thr Leu

35 40 45

Ser Arg Glu Leu Glu Pro Glu Leu Arg Arg Arg Arg Tyr Glu Tyr Asp

50 55 60

His Trp Asp Ala Ala Ile His Gly Phe Arg Glu Thr Glu Lys Ser Arg

65 70 75 80

Trp Ser Glu Ala Ser Arg Ala Ile Leu Gln Arg Val Gln Ala Ala Ala

85 90 95

Phe Gly Pro Gly Gln Thr Leu Leu Ser Ser Val His Val Leu Asp Leu

100 105 110

Glu Ala Arg Gly Tyr Ile Lys Pro His Val Asp Ser Ile Lys Phe Cys

115	120	125
Gly Ala Thr Ile Ala Gly Leu Ser Leu Leu Ser Pro Ser Val Met Arg		
130	135	140
Leu Val His Thr Gln Glu Pro Gly Glu Trp Leu Glu Leu Leu Leu Glu		
145	150	155
Pro Gly Ser Leu Tyr Ile Leu Arg Gly Ser Ala Arg Tyr Asp Phe Ser		
165	170	175
His Glu Ile Leu Arg Asp Glu Glu Ser Phe Phe Gly Glu Arg Arg Ile		
180	185	190
Pro Arg Gly Arg Arg Ile Ser Val Ile Cys Arg Ser Leu Pro Glu Gly		
195	200	205
Met Gly Pro Gly Glu Ser Gly Gln Pro Pro Pro Ala Cys		
210	215	220

<210> 17

<211> 1441

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (95)..(1420)

<400> 17

atattccggt ggctggtctc cggcggcccc gtccccgact gggccccgtg cccccccgcc 60

cccgcggccc cccgccgccg ggccagccgc cacc atg aag aaa ttc ttt cag gag 115

Met Lys Lys Phe Phe Gln Glu

1

5

ttc aag gcc gac atc aag ttc aaa agc gcg gga ccc ggt cag aag ctc 163

Phe Lys Ala Asp Ile Lys Phe Lys Ser Ala Gly Pro Gly Gln Lys Leu

10

15

20

aaa gag tcc gtg ggg gaa aag gcc cac aaa gag aag ccc aac cag cca 211

Lys Glu Ser Val Gly Glu Lys Ala His Lys Glu Lys Pro Asn Gln Pro

25

30

35

gcc ccc agg ccg ccc cgc cag gga ccc acc aat gag gca cag atg gca 259

Ala Pro Arg Pro Pro Arg Gln Gly Pro Thr Asn Glu Ala Gln Met Ala

40

45

50

55

gcc gct gcc gcc cta gcc cgg ctg gag cag aag cag tcc cgg gcc tgg 307

Ala Ala Ala Ala Leu Ala Arg Leu Glu Gln Lys Gln Ser Arg Ala Trp

60

65

70

ggc ccc aca tcg cag gac acc atc cga aac cag gtg aga aag gaa ctt 355

Gly Pro Thr Ser Gln Asp Thr Ile Arg Asn Gln Val Arg Lys Glu Leu

75

80

85

caa gcc gaa gcc acc gtc agc ggg agc ccc gag gcc cca ggg acc aac 403

Gln Ala Glu Ala Thr Val Ser Gly Ser Pro Glu Ala Pro Gly Thr Asn

90

95

100

gtg gta tct gag ccc aga gag gaa ggc tct gcc cac ctg gct gtg cct 451

Val Val Ser Glu Pro Arg Glu Glu Gly Ser Ala His Leu Ala Val Pro

105	110	115	
ggc gtg tac ttc acc tgt ccg ctc act ggg gcc acc ctg agg aag gac 499			
Gly Val Tyr Phe Thr Cys Pro Leu Thr Gly Ala Thr Leu Arg Lys Asp			
120	125	130	135
cag cgg gac gcc tgc atc aag gag gcc att ctc ttg cac ttc tcc acc 547			
Gln Arg Asp Ala Cys Ile Lys Glu Ala Ile Leu Leu His Phe Ser Thr			
	140	145	150
gac cca gtg gcc gcc tcc atc atg aag atc tac acg ttc aac aaa gac 595			
Asp Pro Val Ala Ala Ser Ile Met Lys Ile Tyr Thr Phe Asn Lys Asp			
	155	160	165
cag gac cgg gtg aag ctg ggt gtg gac acc att gcc aag tac ctg gac 643			
Gln Asp Arg Val Lys Leu Gly Val Asp Thr Ile Ala Lys Tyr Leu Asp			
170	175	180	
aac atc cac ctg cac ccc gag gag gag aag tac cgg aag atc aag ctg 691			
Asn Ile His Leu His Pro Glu Glu Glu Lys Tyr Arg Lys Ile Lys Leu			
185	190	195	
cag aac aag gtg ttt cag gag cgc att aac tgc ctg gaa ggg acc cac 739			
Gln Asn Lys Val Phe Gln Glu Arg Ile Asn Cys Leu Glu Gly Thr His			
200	205	210	215
gag ttt ttt gag gcc att ggg ttc cag aag gtg ttg ctt ccc gcc cag 787			
Glu Phe Phe Glu Ala Ile Gly Phe Gln Lys Val Leu Leu Pro Ala Gln			
	220	225	230

gat cag gag gac ccc gag gag ttc tac gtg ctg agc gag acc acc ttg 835
 Asp Gln Glu Asp Pro Glu Glu Phe Tyr Val Leu Ser Glu Thr Thr Leu
 235 240 245

gcc cag ccc cag agc ctg gag agg cac aag gaa cag ctg ctg gct gcg 883
 Ala Gln Pro Gln Ser Leu Glu Arg His Lys Glu Gln Leu Leu Ala Ala
 250 255 260

gag ccc gtg cgc gcc aag ctg gac agg cag cgc cgc gtc ttc cag ccc 931
 Glu Pro Val Arg Ala Lys Leu Asp Arg Gln Arg Arg Val Phe Gln Pro
 265 270 275

tcg ccc ctg gcc tcg cag ttc gaa ctg cct ggg gac ttc ttc aac ctc 979
 Ser Pro Leu Ala Ser Gln Phe Glu Leu Pro Gly Asp Phe Phe Asn Leu
 280 285 290 295

aca gca gag gag atc aag cgg gag cag agg ctc agg tcc gag gcg gtg 1027
 Thr Ala Glu Glu Ile Lys Arg Glu Gln Arg Leu Arg Ser Glu Ala Val
 300 305 310

gag cgg ctg agc gtg ctg cgg acc aag gcc atg cgg gag aag gag gag 1075
 Glu Arg Leu Ser Val Leu Arg Thr Lys Ala Met Arg Glu Lys Glu Glu
 315 320 325

cag cgg ggg ctg cgc aag tac aac tac acg ctg ctg cgc gtg cgc ctc 1123
 Gln Arg Gly Leu Arg Lys Tyr Asn Tyr Thr Leu L u Arg Val Arg Leu
 330 335 340

ccc gat ggc tgc ctc ctg cag ggc act ttc tac gct cgg gag cgg ctg 1171
 Pro Asp Gly Cys Leu Leu Gln Gly Thr Phe Tyr Ala Arg Glu Arg Leu
 345 350 355

ggg gcg gtg tac ggg ttc gtc cgg gag gcc ctg cag agc gac tgg ctg 1219
 Gly Ala Val Tyr Gly Phe Val Arg Glu Ala Leu Gln Ser Asp Trp Leu
 360 365 370 375

cct ttt gag ctg ctg gcc tcg gga ggg cag aag ctg tcc gag gac gag 1267
 Pro Phe Glu Leu Leu Ala Ser Gly Gly Gln Lys Leu Ser Glu Asp Glu
 380 385 390

aac ctg gcc ttg aac gag tgc ggg ctg gtg ccc tct gcc ctc ctg acc 1315
 Asn Leu Ala Leu Asn Glu Cys Gly Leu Val Pro Ser Ala Leu Leu Thr
 395 400 405

ttc tcg tgg gac atg gct gtg ctg gag gac atc aag gcc gcg ggg gcc 1363
 Phe Ser Trp Asp Met Ala Val Leu Glu Asp Ile Lys Ala Ala Gly Ala
 410 415 420

gag ccg gac tcc atc ctg aaa ccc gag ctc ctg tca gcc atc gag aag 1411
 Glu Pro Asp Ser Ile Leu Lys Pro Glu Leu Leu Ser Ala Ile Glu Lys
 425 430 435

ctc ttg tga aataaaagca gggttgcct c 1441
 Leu Leu
 440

<210> 18

<211> 441

<212> PRT

<213> Homo sapiens

<400> 18

Met	Lys	Lys	Phe	Phe	Gln	Glu	Phe	Lys	Ala	Asp	Ile	Lys	Phe	Lys	Ser
1				5					10					15	
Ala	Gly	Pro	Gly	Gln	Lys	Leu	Lys	Glu	Ser	Val	Gly	Glu	Lys	Ala	His
			20					25					30		
Lys	Glu	Lys	Pro	Asn	Gln	Pro	Ala	Pro	Arg	Pro	Pro	Arg	Gln	Gly	Pro
			35				40						45		
Thr	Asn	Glu	Ala	Gln	Met	Ala	Ala	Ala	Ala	Ala	Leu	Ala	Arg	Leu	Glu
		50				55					60				
Gln	Lys	Gln	Ser	Arg	Ala	Trp	Gly	Pro	Thr	Ser	Gln	Asp	Thr	Ile	Arg
65					70					75				80	
Asn	Gln	Val	Arg	Lys	Glu	Leu	Gln	Ala	Glu	Ala	Thr	Val	Ser	Gly	Ser
				85					90					95	
Pro	Glu	Ala	Pro	Gly	Thr	Asn	Val	Val	Ser	Glu	Pro	Arg	Glu	Glu	Gly
			100					105						110	
Ser	Ala	His	Leu	Ala	Val	Pro	Gly	Val	Tyr	Phe	Thr	Cys	Pro	Leu	Thr
			115					120						125	
Gly	Ala	Thr	Leu	Arg	Lys	Asp	Gln	Arg	Asp	Ala	Cys	Ile	Lys	Glu	Ala
			130				135							140	
Ile	Leu	Leu	His	Phe	Ser	Thr	Asp	Pro	Val	Ala	Ala	Ser	Ile	Met	Lys
145					150					155				160	
Ile	Tyr	Thr	Phe	Asn	Lys	Asp	Gln	Asp	Arg	Val	Lys	Leu	Gly	Val	Asp
				165					170					175	
Thr	Ile	Ala	Lys	Tyr	Leu	Asp	Asn	Ile	His	Leu	His	Pro	Glu	Glu	Glu

180	185	190
Lys Tyr Arg Lys Ile Lys Leu Gln Asn Lys Val Phe Gln Glu Arg Ile		
195	200	205
Asn Cys Leu Glu Gly Thr His Glu Phe Phe Glu Ala Ile Gly Phe Gln		
210	215	220
Lys Val Leu Leu Pro Ala Gln Asp Gln Glu Asp Pro Glu Glu Phe Tyr		
225	230	235
Val Leu Ser Glu Thr Thr Leu Ala Gln Pro Gln Ser Leu Glu Arg His		
245	250	255
Lys Glu Gln Leu Leu Ala Ala Glu Pro Val Arg Ala Lys Leu Asp Arg		
260	265	270
Gln Arg Arg Val Phe Gln Pro Ser Pro Leu Ala Ser Gln Phe Glu Leu		
275	280	285
Pro Gly Asp Phe Phe Asn Leu Thr Ala Glu Glu Ile Lys Arg Glu Gln		
290	295	300
Arg Leu Arg Ser Glu Ala Val Glu Arg Leu Ser Val Leu Arg Thr Lys		
305	310	315
Ala Met Arg Glu Lys Glu Glu Gln Arg Gly Leu Arg Lys Tyr Asn Tyr		
325	330	335
Thr Leu Leu Arg Val Arg Leu Pro Asp Gly Cys Leu Leu Gln Gly Thr		
340	345	350
Phe Tyr Ala Arg Glu Arg Leu Gly Ala Val Tyr Gly Phe Val Arg Glu		
355	360	365
Ala Leu Gln Ser Asp Trp Leu Pro Phe Glu Leu Leu Ala Ser Gly Gly		
370	375	380
Gln Lys Leu Ser Glu Asp Glu Asn Leu Ala Leu Asn Glu Cys Gly Leu		
385	390	395
Val Pro Ser Ala Leu Leu Thr Phe Ser Trp Asp Met Ala Val Leu Glu		
405	410	415

Asp Ile Lys Ala Ala Gly Ala Glu Pro Asp Ser Ile Leu Lys Pro Glu

420

425

430

Leu Leu Ser Ala Ile Glu Lys Leu Leu

435

440

【図面の簡単な説明】

【図 1】

クローン HP 0 1 1 2 4 がコードするヒト蛋白質と、ヒトアシル C o A 結合蛋白質のアミノ酸配列を比較した図である。

【図 2】

クローン HP 0 2 2 4 1 がコードするヒト蛋白質と、アフリカツメガエルリボソーム蛋白質 L 2 4 様蛋白質のアミノ酸配列を比較した図である。

【図 3】

クローン HP 1 0 1 0 1 がコードするヒト蛋白質と、線虫仮想蛋白質 C 3 2 E 8. 5 のアミノ酸配列を比較した図である。

【図 4】

クローン HP 1 0 3 7 0 がコードするヒト蛋白質と、ショウジョウバエ仮想蛋白質 CG 1 1 5 3 4 のアミノ酸配列を比較した図である。

【図 5】

クローン HP 1 0 4 2 7 がコードするヒト蛋白質と、線虫仮想蛋白質 Y 1 0 6 G 6 H. 8 のアミノ酸配列を比較した図である。

【図 6】

クローン HP 1 0 5 1 6 がコードするヒト蛋白質と、ショウジョウバエ仮想蛋白質 CG 1 4 1 3 0 のアミノ酸配列を比較した図である。

【図 7】

クローン HP 1 0 5 8 0 がコードするヒト蛋白質と、ショウジョウバエ仮想蛋白質 CG 5 4 6 9 のアミノ酸配列を比較した図である。

特 2 0 0 0 - 1 6 0 8 5 1

【図4】

HP10370 1' MEYDEKLARFRQAHLNPFNKQSGPRQHEQGPGEFVDVTPPEALPELPPGEPEF
DMCG115 1" MSSWKDSLTSIPGTVAQLINESASNLHASTLGSTVGLGSGSTGSGSEAGGSEESGPQ
HP10370 55' RCPERVMDLGLSEDFSRPVGLFLASDVQQLRAIECKQVILELPEQSEKQKDAVRLI
DMCG115 61" GAEYRALPIPASLVREQWRLFTSDANIQDLQAAIAHCRDLVLLSEELSEERRWLVRLV
HP10370 115' HLRKLQELKDPNE-----DEPNIRVLLHRRF-----YKEKSKSVKQTC DK
DMCG115 121" DLRYSLQELEFEAQEQHSLSSDMVVMNAIRAVVGHFVPHPHGKRNRLQAAAKRNYCDH
HP10370 156' CNTIIINGLIQTWYTCGYRCHSKCLNLSKPCVSSKVSHQAEYELNICPETGLDSQDY
DMCG115 181" CTTIISVVQNSYVSCDGLVHQKIDGVKRVCAHVLVSEHQPISEICPEIGLASQGY
HP10370 216' RCAECRAPISLRGVPEARQCDYTGQYCYCHWNDLAVIPARVVHNDFFPRKVSRCSM
DMCG115 241" KCAECGTMLNKNWTWIEPRLCDYSGLYCPCRNWDSNFIPIRIHNWDFSPRRVSRFAL
HP10370 276' RYLALMVSFRPVLRLREINPLLSYVEELVEIRKLQDILLMKPYFITCREAMEARLL-LQ
DMCG115 301" QEIRLFLANKPLIRLEEDNPKLFVTFVEKLCVKKLRQNLVHMRHYLAACKIASLKLVDQQ
HP10370 335' LQDRQHFFVNDENYSVQDLVDVHAGRLGCSLTEIHTLFAKHKLDCERCQAKGFVCELCR
DMCG115 361" LGVRRHLAQSNFYSLSDLSQVESGALSEFLQGVFKAFNDHIR-SCPMCLAQAYICEICS
HP10370 395' EGDVLFPPDSTSVACDSAVFHRDCYTDNSTCPCARLSLRKQSLFQEPGPDVEA
DMCG115 420" NNEVIFPDDGCIKDCQNSIFHRVCLTRKNMICPKIRIQERRLQLDRMKSTEDDDDDDD

【図 5】

```
HP10427 1' MAGPAAAFRRRLGALSGAALGFASYGAH---GAQFPDAYGKELFDKANKHHFTHSLALLG  
... **...***.....**.* ... * .. **..* .....*****.  
CEY106G 1" MSPIIRLAGLSGAVAISIGAYGSHVLRDNPISIDERRRTAFDTASRYHLIHSLLALLA  
  
HP10427 58' VPHCRKPLWAGLLASGTTLCTSFYYQALS G D P S I Q T L A P A G G T L L L L G W L A I A L  
* * * . . * . . . . . * . . . . . * . . . . . * . . . . . *  
CEY106G 57" SPAARFPLVTAGLTAGITLFCGPCIHYSISGVETTRKYTFPIGGVTLIILAWLSFIL
```

【 6】

```

HP10516 1' MAGTGLALRTLPSPWVRGSPSVLSRLQDAVVRPGFLSTAEETLSRELEPELRRR
DMCG141 1" . . . . . * * * * *
MRIITDFISEPEEQQLHEIEPYMSRLR

HP10516 61' YEYDHWDAIHGFRETEKSRWSEASHAILQRVQAAAFPGQTLLSSVHVLDLEARGYIKP
DMCG141 29" ** . . . . . * * * * *
YEYDHWDDAIHGFRETERKKWFPKNREILERVQRVAF--DGAVMPPYVHILD LAPDGVIKP

HP10516 121' HVDSEKFCGATLGLSLLSPVMRLVHT-----QEPG
DMCG141 87" *** . . . . . * * * * *
HVDSTRYCGNTISGISLLSDEVMRLVRTDEQRYQQSSGTATDPNSQGSFPAAYRHQPE

HP10516 153' EWL-----ELLLEPGSLYLILRGSAKYDFSHAILRDEESFFGERRIPRGRRISVICRSLP
DMCG141 147" * . . . . . * * * * *
ASLKNNFYADILLPRRSLYIMSHYARYKFTHEILAKEHSQFQALVPRTRRRIIICRNEP
    
```

【 7 】

HP10580 1' MKKTFQEFKADIKFK
**
DMCG546 61" TLKVLIVLLGTCVLAGYWSIYKVITTEKVFVSTLKEIEELKLSKAEEAFKLNLTGS
HP10580 16' SAGPQOKLESVGEKAHKEPNQ--PAPRPPRQGTNEAOMAAAAALAEQKQSRANQP
* * * * * *
DMCG546 121" GNGTGHKINSFQETPPSSSSRQKYEVPPKRNIEISNEARAAABAALARDKRTSREFN-
HP10580 74' TSQDTIRNQVRNELQAEATVSGSP-EAPGTNVVSEPREGSA-HLAVPGVYFTCPL-TGA
** ** ** ** ** **
DMCG546 180" TSLSAVTAQAKRELEAERRQREEMGTPTSTSTSTASGGDTNLTACGVTFRCPLEEE
HP10580 131' TLRKQORDACIKEAILLHFSTDPVAASIMKIYTNKQDQDRVKLGVDVTIAKYLDNIHLHPE
* * * * * *
DMCG546 240" ILPKSVWKVRIKEFLYQOLEADRGLTACLI IHNCN-VKEKADECIATLIRVLENLINKPE
HP10580 191' EEKYRKIKLQNKVVFQERINCLEGTHTHFFEAIGFQKVLPAQDQEDDPEEFYVLSETTLAQ
*** ** ** ** ** **
DMCG546 299" EEKFKIRMSNKIFSEKRVYVEGALDVLOAAGFNEV-----QIDGEPFLMTKEQTERD
HP10580 251' QSLERHKEQLLAAEFVRAKLDQRKRVQPSPLASQFELPGDFFNLTAEEIKREQLRSEA
* * * * * *
DMCG546 353" LDLEFLVEALNSSEIIPLELDNRINKVLLFSQ-ACRVALLPEFYRLSPSEIKKEQQLRSEA
HP10580 311' VERLSVLRTRAMREKEQGLKKNYTLRLVRLPDGCLLQGTTFYARERLGAIVYGVFREAL
... .. * * * * * *
DMCG546 412" IAQSQMLRTKAMRERESQRLRMRYALVRVFPNGLFIOGTFFNVYKISDVFFSVQSCIL
HP10580 371' QSDWLPFELLASGGQKLSHDE-NLALNECGLVPSALLTFSDMAVLEDIRAAGAEPDSIL
... * * * * * *
DMCG546 472" ADESLDFSLVNSDGLGDEHLEKTYCKLIPTNLTLLFSANDTPAPLQTDINYLYKEDLL
HP10580 430' KPBLISATIEKLL
DMCG546 532" MLVQAM

【書類名】 要約書

【要約】

【課題】 精製ヒト蛋白質、この蛋白質をコードしている完全長 cDNA を含む DNA 断片、この DNA 断片の発現ベクター、この発現ベクターによる形質転換細胞およびこの蛋白質に対する抗体を提供する。

【解決手段】 配列番号 2、4、6、8、10、12、14、16 または 18 のいずれかのアミノ酸配列を有する精製ヒト蛋白質、配列番号 1、3、5、7、9、11、13、15 または 17 の翻訳領域の塩基配列を有する DNA 断片、この DNA 断片の発現ベクター、この発現ベクターによる形質転換細胞、およびこの蛋白質に対する抗体。

【選択図】 なし

特 2 0 0 0 - 1 6 0 8 5 1

出 願 人 履 歴 情 報

識別番号 [3 9 6 0 2 0 8 0 0]

1. 変更年月日	1 9 9 8 年 2 月 2 4 日
[変更理由]	名称変更
住 所	埼玉県川口市本町4丁目1番8号
氏 名	科学技術振興事業団